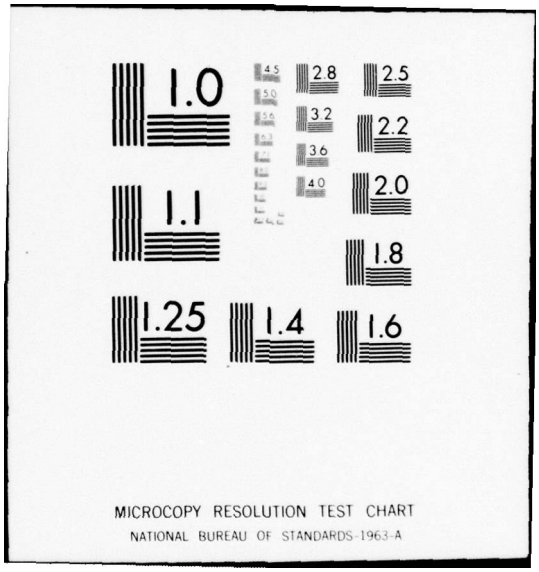
AD-A051 259 GEORGIA INST OF TECH ATLANTA F/G 12/1
A COST OPTIMAL APPROACH TO SELECTING A FRACTIONAL FACTORIAL DES--ETC(U)
JUN 77 W F FRIESE

UNCLASSIFIED NL

1 OF 1
AD
A051259

END
DATE
FILMED
4-78
DDC


1.0
1.1
1.25
1.4
1.6
1.8
2.0
2.2
2.5
2.8
3.2
3.6
4.0
4.5
5.0
5.6
MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

A COST OPTIMAL APPROACH TO SELECTING A FRACTIONAL FACTORIAL DESIGN

A THESIS

Presented to

The Faculty of the Division of Graduate Studies

By

William Francis Friese, Jr.

In Partial Fulfillment

of the Requirements for the Degree

Master of Science

in Operations Research

Georgia Institute of Technology

June, 1977

(1)

(6) A COST OPTIMAL APPROACH TO SELECTING A FRACTIONAL FACTORIAL DESIGN.

(9) Master's thesis,

A THESIS

Presented to

The Faculty of the Division of Graduate Studies

By

(10) William Francis Friese, Jr.

(11) Jun 77

(12) 93 p.

In Partial Fulfillment

of the Requirements for the Degree

Master of Science

in Operations Research

Georgia Institute of Technology

June, 1977

A COST OPTIMAL APPROACH TO SELECTING

A FRACTIONAL FACTORIAL DESIGN

Approved:

Douglas C. Montgomery

Dr. Douglas C. Montgomery, Chairman

Russell G. Heikes

Dr. Russell G. Heikes

Ronald L. Rardin

Dr. Ronald L. Rardin

Date approved by Chairman: 5-13-77

ACKNOWLEDGMENTS

I would like to express my appreciation to everyone who has assisted me in the conduct of this research. In particular, I thank Dr. Douglas C. Montgomery, the chairman of my thesis committee, for his assistance in formulating the problem and his guidance throughout the research, and the other members of my thesis committee, Dr. Russell G. Heikes and Dr. Ronald L. Rardin, for their suggestions and constructive evaluations.

I would also like to thank the United States Army for the opportunity to attend graduate school and for the assistance received in the conduct of this research.

Finally, I would like to express special thanks to my wife and daughters for the support and understanding during this research.

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

LIST OF TABLES

LIST OF ILLUSTRATIONS

SUMMARY

This research develops a multistage decision process designed to obtain the maximum amount of information from the evaluation of a factorial design while minimizing the amount of resources used in obtaining the information. The use of screening experiments in building the factorial design is investigated in order to maximize the amount of information gained. The use of sequential analysis procedures to terminate experimentation at the earliest possible time is investigated in order to minimize the amount of resources used. The research is limited to 2^n factorial designs involving univariate response models assumed to come from a normal population; however, the procedure can be easily extended to any factorial design.

The approach is demonstrated for an operational test involving a 2^6 factorial design and the results are compared to "classical" procedures. The sensitivity of the required input parameters is investigated and related applications are discussed.

The proposed approach is found to be a viable method of designing, conducting, and evaluating an operational test involving a factorial experiment.

CHAPTER I

INTRODUCTION

Background

The goal of any type of experimentation is to obtain information about the system under investigation. Naturally, the more information that can be obtained the better. Unfortunately, obtaining more information often requires increased experimentation and the resources available may become a limiting factor. The shortage of any necessary resource, whether it be time, money or materials can greatly hinder the conduct of the experiment and even preclude obtaining the desired amount of information from the experimentation.

Much effort has been devoted to the problem of how to best utilize the experimental resources available in order to gain the most information from them. The development of systematic experiments such as factorial experiments and procedures for fractional factorial experiments have done much to improve the use of available resources. Screening experiments have aided by eliminating needless experimentation involving unimportant factors. Also, sequential experiments have helped to obtain the desired information from less resources.

The problem of gaining the most information from limited resources can be found everywhere. From agriculture to industry to the defense establishment, everyone is interested in getting the most for their money.

Material Acquisition Process

The requirement to maintain a modern, well-equipped Army requires constant evaluation and reevaluation of existing Army equipment to insure that it is adequate to fulfill operational requirements. Based on the ever changing nature of these requirements, the Army is involved in a continuous process of upgrading its current equipment and procuring new items of equipment. Basically, the Army satisfies its needs for new equipment in three ways: buying equipment already developed, evolutionary development of current standard equipment, and initiation of new material development programs. All of these procurement methods can be extremely costly in terms of time, material, and money. As a result, both the Department of Defense and the Department of the Army have highly structured material procurement policies [7, 18] whose objectives are to minimize the costs incurred in acquiring material systems while insuring that the performance of those systems is adequate to meet operational requirements.

Once the requirement for a new or updated system has been formalized, the proposed system will go through three phases of development: conceptual development, validation, and full scale development, before the system receives approval for full production and purchase by the Department of Defense.

At the end of each phase, the Defense Systems Acquisition Review Council (DSARC) meets to provide information and recommendations to the Secretary of Defense. Based on these recommendations, the Secretary of Defense may decide to cancel further consideration of the system, require further system evaluation prior to proceeding on to the next stage, or permit the system to pass on to the next stage of development.

At Department of the Army level, there is a similar advisory body, the Army Systems Acquisition Review Council (ASARC), whose principal function is to provide the DSARC with the Army's recommendations concerning the item of equipment in question.

Testing

To aid the ASARC in its recommendations, testing is conducted to demonstrate how well the material system meets its technical and operational requirements; provide data to assess developmental and operational risks for decision making; verify that the technical, operational, and support problems identified in previous testing have been corrected; and to insure that all critical issues to be resolved by testing have been adequately considered. Two types of testing, Developmental Testing (DT) and Operational Testing (OT) are conducted. DT is conducted to demonstrate that the engineering design and development process is complete, that design risks have been minimized, and that the system will meet required specifications. It is performed by the material developer who then forwards the results to the ASARC.

OT is conducted to estimate the system's military utility, operational effectiveness, operational suitability, and the need for any modifications. OT can also provide data on organization, personnel requirements, doctrine and tactics for the new system. OT is performed under the supervision of the U.S. Army Operational Test and Evaluation Agency (OTEA) by operational and support personnel of the type and qualifications expected to use and maintain the system once it is deployed.

As a safeguard and as a further validation measure, DT and OT test

designs are prepared and the test results are evaluated independently. The actual testing of the item of equipment under DT and OT may, however, be conducted concurrently to reduce delays and system acquisition costs. The relationship between the conduct of the different phases of testing and the meeting of the Acquisition Review Councils is shown in Figure 1.

Operational Testing

OTEA serves the following major functions:

A. Insures user testing is effectively planned, conducted, and evaluated with emphasis on adequacy, quality, and credibility of all user testing.

B. Actively participate in the conduct of and provide independent evaluations of operational tests conducted on selected items of equipment.

C. Develop and recommend policy on user testing.

D. Develop and promulgate user test and evaluation methodology.

E. Develop measures of effectiveness and provide estimates on amount of resources (sample size) necessary to detect differences in military utility, operational effectiveness, and operational suitability with a specified confidence level.

Generally three phases of Operational Testing are conducted, one phase prior to each meeting of the ASARC. OT will compare the performance characteristics of the new system against the current system if it is designed as a replacement item, against a higher level system if it is designed as a component of some larger system, or against a set of performance standards if the item is totally new to the Army inventory. Due to the nature of the items being tested, such as missile systems or high value items that require destructive type testing, OTEA is often limited in the number of experimental runs that can be made on an item of equipment. As

PHASE	REQUIREMENT DETERMINATION	CONCEPTUAL DEVELOPMENT	VALIDATION	FULL SCALE DEVELOPMENT	PRODUCTION & DEPLOYMENT
SYSTEM STATUS	CONCEPTUAL	EXPERIMENTAL PROTOTYPE	ENGINEERING DEVELOPMENT PROTOTYPE	PRODUCTION PROTOTYPE	FINAL MODEL
ACQUISITION PROCESS	ROC	OT I ASARC I DSARC I DT I	OT II ASARC II DSARC II DT II	OT III ASARC III DSARC III DT III	

Figure 1. Major Defense Systems Acquisition Process [51].

a result, OTEA is very much interested in developing a methodology for designing, planning, and evaluating operational tests for a limited sample size while, at the same time, maximizing the information gained from the test.

Problem, Objective, Scope

This problem was motivated by a task requirement presented by OTEA:

> Obtaining maximum information from minimum sample size is an inherent and recurring problem in operational testing and has significant impact on final evaluations. In designing and conducting operational tests, resource restrictions often dictate a test having an extremely small sample size but a number of influencing factors of two or more conditions. This results in a relatively large number of combinations, especially considering the number of observations to be obtained. This study is to investigate the feasibility of developing a method for maximizing the information gained from a test.

A related problem, that of designing, planning, and evaluating operational tests of limited sample size has already been addressed by Russ (48). In his research, he developed an algorithm for determining the optimum constrained sample size for a full factorial experiment based on a specified amount of information required by the test evaluator. The objective of this study will be to develop a sequential method of designing and conducting an operational test in order to gain equivalent information from an even smaller sample size.

The scope of this research will be limited in the following areas:

A. All factors in the factorial design will appear at only two levels and will be considered as fixed factors. The extension of the problem to cases where the factors appear at other than two levels will pose no problems for anyone familiar with the analysis of factorial

experiments.

B. Only univariate response models are considered.

C. The measured response is assumed to come from a normal population with mean, μ, and variance, σ^2. This would appear to be a valid assumption in light of the frequency with which it occurs in everyday situations.

D. The hypotheses to be tested are assumed to be of the form H_0: $\mu=\mu_1$ vs. H_1: $\mu - \mu_1 > d$, This assumption is based on the fact that Operational Testing is generally performed to test the performance characteristics of one system against another system or against a set of standards. The decision to accept or reject the new system is based on whether or not the performance characteristics of the new system exceed the old system by a specified margin, d, or not.

The research will consist of a review of full and fractional factorial design construction and analysis and sequential analysis methodology. A proposed sequential method for building a factorial design is then developed and applied to a previously performed operational test to demonstrate its use in obtaining the same information from a reduced sample size.

CHAPTER II

REVIEW OF OTHER APPROACHES AND RELATED TECHNIQUES

This chapter will present a review of the construction and analysis of a factorial experiment and will review other approaches to the sequential analysis problem. Both the case where the variance is known and the case where the variance is unknown are examined in the sequential analysis review.

Factorial Experiments

Full Factorials

A factorial experiment is the term used to denote the experimental design in which all levels of a given factor are combined with all levels of every other factor in the experiment. In the case where each of the n factors can be measured at the same number of levels, say m, the experiment is referred to as an m^n factorial. For example, in a 2^n factorial, each of the n factors would appear at two levels so there are 2^n different treatment combinations. In a 2^n, the levels of the factors are arbitrarily denoted as the high and low levels. Generally accepted notation is to represent the various factors using capital letters, A, B, C, etc., and to denote the various treatment combinations using lower case letters, a, b, c, etc. The letters present in the treatment combination indicate those factors appearing at their high level and the letters absent indicate those factors appearing at their low level. For example, the treatment combination a denotes factor A at its high level and all other

factors at their low levels and cd denotes factors C and D at their high levels and all other factors present in the experiment at their low levels. The notation (1) is used to represent the treatment combination where all factors appear at their low levels.

The common representation for the univariate response model for the 2^2 factorial design is:

$$y_{ijk} = \mu + \alpha_i + \beta_j + \alpha_{ij} + \varepsilon_{ijk} \qquad (2\text{-}1)$$

where

y_{ijk} = the k^{th} observation at the i^{th} level of factor A and the j^{th} level of factor B.

μ = the population mean of the observations.

α_i = effect of factor A at level i, i=1,2.

β_j = effect of factor B at level j, j=1,2.

$\alpha_i\beta_j$ = interaction effect of factor A at level i and factor B at level j.

ε_{ijk} = error associated with observation y_{ijk}.

The restriction $\Sigma\Sigma_{ij} n_{ij}\alpha_i = \Sigma\Sigma_{ij} n_{ij}\beta_j = 0$ where n_{ij} number of observations in the ij^{th} cell also applies. This model can easily be extended to n factors by the addition of appropriate terms.

In the case where $n_{ij} = 1$ for all i and j, there is no separate estimate for error and, in general, higher order interactions are assumed to be negligible and are pooled to estimate the error. The most desirable situation is the case where there is more than one observation per cell and where the number of observations in each cell is the same.

The effect due to a factor is defined to be the change in the response produced by a change in the level of that factor. Interaction between factors exist when a change in one factor produces a different change in the response variable at one level of another factor than at the other level of that factor.

All of the treatment combinations present in a factorial experiment can be expressed in standard order. The simplest method of writing the treatment combinations in standard order is the method of signs as demonstrated in Montgomery (42). This method consists of listing all n factors as column headings. The first column consists of a total of 2^n alternating minus and plus signs starting with a minus sign. The second column consists of alternating pairs of minus and plus signs starting with two minus signs. The third column consists of alternating sets of four minus and plus signs. In general, the column for the n^{th} factor will consist of alternating sets of 2^{n-1} minus and plus signs always beginning with a set of minus signs until there are a total of 2^n entries in that column. The combinations of minus and plus signs in each row are the treatment combination represented by that row where the minus sign indicates a factor at its low level and a plus sign indicates a factor at its high level. This procedure is illustrated for a 2^4 factorial experiment in Table 1.

Blocking

In experimental designs involving a large number of trials, it may be impossible to perform the entire experiment under homogeneous conditions. In this case, the experimenter may have to perform the experiment in blocks. In order to separate the full factorial into

Table 1. Method of Signs for a 2^4 Factorial

Factor A	Factor B	Factor C	Factor D	Treatment Combination
-	-	-	-	(1)
+	-	-	-	a
-	+	-	-	b
+	+	-	-	ab
-	-	+	-	c
+	-	+	-	ac
-	+	+	-	bc
+	+	+	-	abc
-	-	-	+	d
+	-	-	+	ad
-	+	-	+	bd
+	+	-	+	abd
-	-	+	+	cd
+	-	+	+	acd
-	+	+	+	bcd
+	+	+	+	abcd

blocks, the experimenter must decide which effect (s) he is not interested in or can assume to be negligible. Usually, this is the highest order interaction or interactions present. Once this decision is made, a defining contrast, an expression stating which effects are to be confounded (not estimable) with blocks, is established. Hicks [25] and Kempthorne [36] demonstrate several methods for separating the treatment combinations into blocks. Each block should contain a number of treatment combinations no greater than the number of experimental trials that can be performed at one time.

Regular Fractional Factorials

In the case of a large number of treatment combinations present in an experiment, it may not be either physically or economically possible for the experimenter to obtain an observation at every treatment combination. However, it is possible to run only a fraction of the experiments and still obtain the same information as if the entire experiment had been performed. The general procedure is known as fractional replication.

Box and Hunter [29] list the following as major uses of fractional factorial designs:

A. When certain interactions can be assumed non-significant from prior knowledge, these interactions can be used as generators in the separation of treatment combinations into their respective fractions in such a way as to make efficient use of the analysis of the fractional design.

B. In screening experiments where it is expected that the effects of all but a few of the variables studied will be negligible, the use of fractional factorials as a screening experiment will enable the

experimenter to determine which factors are important and which are not without performing costly and unnecessary experimentation. The insignificant factors may then be set at standard levels and the significant factors explored more comprehensively during further experimentation using a smaller, less costly factorial design.

c. Where groups of experiments are run in sequence and ambiguities remaining at a given stage of experimentation can be resolved by later groups of experiments, the full factorial can be fractionated so that the fractional experiment performed yields information on only the factors in question.

d. Where certain major variables, which may interact, are to be studied simultaneously with other minor variables whose influence, if any, can be described by main effects only, the fractional design can be established so as to confound the major variable main effects with only two factor interactions of the minor variables but these interactions are assumed to be negligible so they have no effect on the main factors.

In order to separate the full factorial into fractions, it is necessary for the experimenter to determine which effects he wishes an estimate for and which effects he is willing to assume as negligible. Fractional factorial designs can be classified in the following manner for convenience:

a. Resolution II designs: main effects are not confounded with each other but are confounded with two factor interactions and two factor interactions are confounded with each other. The smallest possible resolution III design is the 2^{3-1} design which can be generated

using either I = ABC or I = -ABC as the defining relationship. This design can be easily generated by the method of signs. First, a full 2^2 design using - and + signs is written down. The signs in the third column are generated as the product of the signs in the first two columns multiplied by +1 if I = ABC is the generator used and by -1 if I = -ABC is the generator used. This is equivalent to equating the third factor, c, to the product of the first two (C = AB) or to the negative product of the first two (C = -AB). This procedure is illustrated in Table 2.

Table 2. 2^{3-1} Resolution III Design

Full 2^2							
			A	B			
			-	-			
			+	-			
			-	+			
			+	+			
A	B	C	Treatment Combination	A	B	C	Treatment Combination
-	-	+	c	-	-	-	(1)
+	-	-	a	+	-	+	ac
-	+	-	b	-	+	+	bc
+	+	+	abc	+	+	-	ab

b. Resolution IV designs: main effects are not confounded with each other or with two factor interactions but two factor interactions are confounded with each other. The smallest resolution IV design

is the 2^{4-1} which can be generated in the same manner as in the previous example by starting with a full 2^3. A 2^{6-2} Resolution IV design with I = ABCE and I = ABDF as the generators is shown in Table 3.

Table 3. 2^{6-2}_{IV} Design, I = ABCE and I = ABDF

Factor A	Factor B	Factor C	Factor D	E=ABC	F+ABD	Treatment Combination
-	-	-	-	-	-	(1)
+	-	-	-	+	+	aef
-	+	-	-	+	+	bef
+	+	-	-	-	-	ab
-	-	+	-	+	-	ce
+	-	+	-	-	+	acf
-	+	+	-	-	+	bcf
+	+	+	-	+	-	abce
-	-	-	+	-	+	df
+	-	-	+	+	-	ade
-	+	-	+	+	-	bde
+	+	-	+	-	+	abdf
-	-	+	+	+	+	cdef
+	-	+	+	-	-	acd
-	+	+	+	-	-	bcd
+	+	+	+	+	+	abcdef

c. Resolution V designs: no main effects or two factor interaction is confounded with any other main effect or two factor interaction, but two factor interactions are confounded with three factor interactions. The smallest Resolution V design is the 2^{5-1} which can also be constructed in the same manner as the Resolution III and IV designs previously described.

In general, the resolution of a design is equal to the smallest number of characters in any word appearing in the defining relation. The words in the defining relation consist of the generators initially chosen and all of their generalized interactions (products mod 2 on exponents). For the design shown in Table 3, the complete defining relation is $I = ABCE = ABDF = CDEF$.

One problem that will arise in fractionating a full factorial is that two or more effects may have the same numerical value. In this case, the effects are known as aliases and the experimenter must be sure that factors believed to be significant are not aliased with each other. The aliases of any factor can be generated by multiplying that factor by all of the words in the defining relation, mod 2 on exponents. The alias structure for the 2^{6-2} Resolution IV design with defining relation $I = ABCE = ABDF = CDEF$ is shown in Table 4. For a 2^{n-p} fractional factorial design, each effect will have $2^{p}-1$ aliases.

When estimating the effect of a factor for a fractional factorial design, we are really estimating the effect due to that factor and all of its aliases. Therefore, the effect of A is really a measure of the effect due to $A + BCE + BDF + ACDEF$. If, as in the general case, interactions of third order or higher are assumed to be negligible, then

Table 4. Alias Structure for 2^{6-2}_{IV} Design,

I = ABCE = ABDF = CDEF

I	ABCE	ABDF	CDEF
A	BCE	BDF	ACDEF
B	ACE	ADF	BCDEF
C	ABE	ABCDF	DEF
D	ABCDE	ABF	CEF
E	ABC	ABDEF	CDF
F	ABCEF	ABD	CDE
AB	CE	DF	ABCDEF
AC	BE	BCDF	ADEF
AD	BCDE	BF	ACEF
AE	BC	BDEF	ACDF
AF	BCEF	BD	ACDE
CD	ABDE	ABCF	EF
CF	ABEF	ABCD	DE
ACD	BDE	BCF	AEF
ADE	BCD	BEF	ACF

this response will provide a good estimate of the effect due to factor A. The analysis of this effect will be discussed in a later section.

If, as in the case of a large number of factors present in the experiment, a one-half fraction still leaves too many observations to be taken, it is possible to use a smaller fraction. To run a 2^{n-P} fractional factorial design requires the choice of P independent generators

(no chosen generator is a generalized interaction of the others).

Irregular Fractional Factorials

In many cases, especially in an experiment involving a large number of variables, the running of a regular fraction of a 2^n, a 2^{n-P} say, may not be economical. For example, suppose the experiment in question is a 2^7 and the experimenter can afford to perform 50 trials of the experiment. If he wishes clear estimates of all main effects and two factor interaction, a 1/2 replicate of the 2^7 will provide them but it requires 64 trials, more than the allotted number. A 1/4 replicate of the 2^7 requires only 32 trials but "there does not exist a 1/4 replicate of the 2^7 experiment which allows uncorrelated estimates of all main effects and two factor interactions." (1) The logical question that arises is whether or not a plan can be constructed using close to but not more than 50 trials that will yield clear estimates of the main effects and two factor interactions. Addelman [1] and John [33] both propose slightly different solutions to the development of this design.

Addelman defines his irregular fraction as a $K/2^P$ fraction of a 2^n factorial. He builds his irregular fraction by combining the treatment combinations in K distinct $1/2^P$ replicates of the 2^n. In Addelman's irregular fraction plan, no main effect or interaction need be completely confounded with the mean if $K \geq P + 1$ but if $K = P - \mu$, $\mu = 0, 1, 2, \ldots,$ then $u + 1$ effects or interactions and their generalized interactions will be completely confounded with the mean. If this should happen, however, most of the time it will be possible to construct the irregular fraction so that the effects which are completely confounded with the mean will contain at least five factors. In the instances where this

irregular fraction cannot be constructed, some two factor interactions will not be estimable. Addelman shows that the yield of the treatment combinations in his irregular fraction can be expressed in terms of the main effects and two factor interactions in the following manner:

$$y_{ijk\ldots} = \mu \pm \tfrac{1}{2}A \pm \tfrac{1}{2}B \pm \tfrac{1}{2}AB \pm \tfrac{1}{2}C \pm \tfrac{1}{2}AC, \text{ etc. + error} \qquad (2.2)$$

where the sign

on A is - if i=0 and + if i=1

on B is - if j=0 and + if j=1

on AB is - if the product of the signs on A and B is - and is + if the product of the signs on A and B is +, and so on.

In this case, 0 and 1 would indicate the presence of that factor at its low and high levels in the treatment combination in question. In the appendix to his paper, Addelman gives several common and useful irregular fractions with the identity relationship and assumptions required to generate them. One such example is the 3/8 fraction of the 2^7 which would be useful in the situation described at the beginning of this section. Using the identity relationship I = ABCDE = ABF = CDEF = AEG = BCDG = BEFG = ACDFG, the 2^7 can be fractionated into eight blocks of 16 runs each. The experimenter could then pick three of these blocks, yielding 48 experimental trials.

John approaches the problem of irregular fractions by subtracting treatment combinations from the full factorial or by adding treatment combinations to a 1/2 replicate of the full factorial. He defines his designs in terms of the relationship used to generate the missing fraction. John illustrates how the combination of fractions in a certain manner

will form overlapping fractions. Estimates for the effects due to various treatment combinations are then obtained by combining the estimates obtained from the appropriate overlapping fractions.

Analysis of Factorial Experiments

Once the experiment has been performed, the results must be analyzed to determine the significance or insignificance of the various factors and interactions or to perform tests on the various hypotheses concerning the factors. The most commonly used method of analysis is the analysis of variance (ANOVA).

ANOVA makes use of the fact that the ratio of two chi-square random variables divided by their respective degrees of freedom follows an F distribution. Hicks [25] and Hines and Montgomery [26] provide the general format for an ANOVA table for a 2^n factorial experiment with r replications per cell shown in Table 5. The F_0 statistic is found by forming the ratio of ms for the effect in question to ms_E. Comparing this value to the value of the F statistic with 1 and $2^n(r-1)$ degrees of freedom will give a test on the significance of an effect.

The easiest method for computing the sum of squares for a 2^n factorial is the Yates Method [57] which consists of arranging the treatment combinations in standard order and then adding and subtracting the observed response values in pairs a total of n times to obtain an estimate of the contrast due to a treatment combination. The effect due to that treatment combination is then obtained from

$$\text{EFFECT} = (2)(\text{CONTRAST})/r.2^n \qquad (2.3)$$

where r = number of replications per cell. This will usually be one for

Table 5. ANOVA for 2^n with r Replications Per Cell

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F_0
Main Effects--A	SS_A	1	$SS_{A/1}$	
B	SS_B	1	$SS_{B/1}$	
C	SS_C	1	$SS_{C/1}$	
...		...		
2 Factor Interactions--AB	SS_{AB}	1	$SS_{AB/1}$	
AC	SS_{AC}	1	$SS_{AC/1}$	
BC	SS_{AC}	1	$SS_{AC/1}$	
...		...		
3 Factor Interactions--ABC	SS_{ABC}	1		
ABD	SS_{ABD}	1		
BCD	SS_{BCD}	1		
...		...		
4 Factor Interactions, Etc.				
Sum of all Treatment Combinations	SS	2^n-1		
Residual or Error	SS^T	$2^n(r-1)$	$\frac{SS_E}{2^n(r-1)}$	
Total	SS_T	$r(2^n)-1$		

n

$c(n,2) = \frac{n(n-1)}{2}$

$c(n,3) = \frac{n(n-1)(n-2)}{6}$

purposes of this paper. The sum of squares for any effect is then obtained from

$$SS = (CONTRAST)^2/r \cdot 2^n \qquad (2.4)$$

Because of its simplicity, the Yates method is easily programmed for computer application.

The Yates method can also be used in obtaining the sums of squares from a 2^{n-P} fractional factorial by considering the data as having come from a full factorial in n-P variables. The treatment combinations for the full factorial are written in standard order and a letter or letters is added in parentheses to the end of these treatment combinations to represent the treatment combinations actually run. The effect estimated in this manner will then be the effect associated with the treatment combination shown plus all of its aliases. An example of the Yates method for a 2^{4-1} with I = ABCD as the defining relationship is shown in Table 6.

The values in column (1) are obtained by first adding the responses in pairs and then subtracting them in pairs. For example, 82 = 74 + 108, 222 = 92 + 130, 173 = 68 + 105, 228 = 95 + 133 and 34 = 108 - 74, 38 = 130 - 92, 37 = 105 - 68 and 38 = 133 - 95. The values in columns (2) and (3) are obtained by performing the same operations on the values in columns (1) and (2) respectively. It is necessary to perform these operations three times since, in this case, n - P = 4 - 1 = 3.

Table 6. Yates Algorithm for 2^{4-1} with I = ABCD [42]

Treatment Combination	Response	(1)	(2)	(3)	Effect	Estimate of Effect $2\times(3)/2^n$	SS $(3)^2/2^n$
(1)	74	182	404	805	-		
a(d)	108	222	401	147	A+BCD	36.75	2701.125
b(d)	92	173	72	95	B+ACD	23.75	1128.125
ab	130	228	75	5	AB+CD	1.25	3.125
c(d)	68	34	40	-3	C+ABD	-0.75	1.125
ac	105	38	55	3	AC+BD	0.75	1.125
bc	95	37	4	15	BC+AD	3.75	28.125
abc(d)	133	38	1	-3	ABC+D	-0.75	1.125

Sequential Procedures

Introduction

Until recently, commonly accepted statistical procedures involved presenting the data from an already conducted experiment to the statistician and expecting him to provide reasonable conclusions based on an analysis of the data. Seldom was the statistician consulted concerning the methods of collecting the data either prior to or during the experimentation. Lately, however, much more emphasis has been placed on obtaining the advice of a statistician prior to the performance of an experiment and during the actual running of the experiment. This has allowed the statistician to play an important role in designing and monitoring the conduct of experiments. In addition, more emphasis has been

placed on obtaining meaningful results from a reduced number of trials of an experiment due to the rapidly increasing costs associated with experimentation. Fractionating a factorial design is one way of accomplishing this goal. A second useful method is through sequential analysis procedures.

Sequential analysis can best be described as "any statistical procedure in which the final pattern (including the number) of observations is not determined a priori but depends, in some way or other, on the values observed in the course of the work" [34]. During the course of an experiment testing one hypothesis against another, the results are examined after each trial. A decision is then made to either accept one or the other of the hypotheses or to continue sampling based on the results to date. The decision making process of sequential analysis will require three basic rules to define the procedure:

A. The stopping rule will let the experimenter know when the experimentation may be terminated.

B. The terminal decision rule will let the experimenter know which hypothesis to decide in favor of once experimentation has been terminated.

C. The experimentation rule will let the experimenter known which experiment should be performed next should he still be required to continue experimentation.

Curtailed Sampling

Although Wald [52] is generally credited with originating sequential analysis procedures in the mid 1940's, there were some heuristic approaches made to this area prior to Wald. One such approach is known as curtailed sampling and is used in determining whether to accept or

as curtailed sampling and is used in determining whether to accept or reject a lot of some item. "Classical" sampling procedures would call for a fixed sample of size n to be inspected and the decision to accept or reject the entire lot would be based on the number of defective items found in that sample. Curtailed sampling procedures would stop sampling as soon as it was obvious which decision would have to be made. For example, if the fixed size test called for rejecting a lot if more than two defective items were found in a sample of size ten, curtailed sampling would call for termination of sampling as soon as:

A. Three defective items were found,

B. Nine items had been inspected and only one defective item was found, or

C. Eight items had been inspected and no defective items were found.

Tests on the Mean of a Normal Distribution with Known Variance

Sequential Probability Ratio Test. Wald's Sequential Probability Ratio Test (SPRT) is designed to test one simple hypothesis against another. If a random variable, x, has distribution $f(x,\theta)$, the SPRT will test $H_0:\theta = \theta_0$ vs. $H_1: \theta = \theta_1$ and will decide in favor of either H_0 or H_1 based on a series of observations of $x(x_1, x_2, \ldots, x_n)$ and on preselected probabilities of type I (reject H_0 when it is true) and type II (accept H_0 when it is false) errors. If these probabilities are denoted as α and β respectively, there are four possible outcomes as shown in Table 7.

For a fixed sample size, n, Neyman and Pearson, as explained in Hoel, Port and Stone [28], have shown that the most powerful test (that test giving the smallest β) depends on the likelihood ratio:

$$\frac{L_1(x)}{L_0(x)} = \frac{\prod_{i=1}^{n} f(x_i, \theta_1)}{\prod_{i=1}^{n} f(x_i, \theta_0)} \tag{2.5}$$

where $f(x_i, \theta_i)$ = probability of observing x_i given that θ_i is true. This test will decide in favor of H_0 if $L_1(x)/L_0(x)$ is less than some constant, k, and will decide against H_0 if $L_1(x)/L_0(x)$ is greater than k. The value of k can be chosen to insure the desired α value for the test and the number of observations, n, can be chosen to insure the desired power, $1 - \beta$, for the test.

Table 7. Outcomes and Probabilities of SPRT

Outcome	Required Probability
θ_0 is true and test decides in favor of θ_0	$1 - \alpha$
θ_0 is true and test decides in favor of θ_1	α
θ_1 is true and test decides in favor of θ_1	$1 - \beta$
θ_1 is true and test decides in favor of θ_0	β

The SPRT is similar to this procedure and incorporates the following decision rules:

A. Stop sampling and decide in favor of H_0 as soon as $L_1(x)/L_0(x)$ is less than some constant, B.

B. Stop sampling and decide in favor of H_1 as soon as $L_1(x)/L_0(x)$ is greater than some constant, $A > B$.

C. Continue sampling as long as $B < L_1(x)/L_0(x) < A$.

Wald [52] has proven that good approximations for the constants A and B are:

$$A \cong (1-\beta)/\alpha \tag{2.6}$$

$$B \cong \beta/(1-\alpha) \tag{2.7}$$

Since α and β are generally taken to be less than .5, the following relationship will hold:

$$0 < B < 1 < A \tag{2.8}$$

The Sequential Probability Ratio Test possesses many properties that make it very useful when testing a simple hypothesis against a simple alternative.

A. Wald and Wolfowitz [53] have shown that for all sequential tests having the same α and β probabilities of Type I and Type II error, the SPRT will require the fewest number of observations when θ has true value equal to θ_0 or θ_1.

B. Should the true value of θ lie somewhere between θ_0 and θ_1, Wetherill [56] has shown that the expected number of observations may be much larger than the fixed sample size plan with the same α and β errors.

C. Wald has proposed truncating the SPRT at some fixed value, n_0, should this occur. This will change the probabilities of Type I and Type II errors so the problem is to make n_0 large enough so as to have a negligible effect on these probabilities. By denoting $P_i(n_0)$ as the probability of a sample of size n_0 rejecting H_i under the truncated process and accepting H_i under the non-truncated process, Wald derives

upper bounds for these two probabilities and uses them to determine upper bounds for the new α and β errors respectively. The upper bounds for $P_0(n_0)$ and $P_1(n_0)$ are

$$\bar{P}_0(n_0) = G(u_2) - G(u_1) \quad \text{and} \quad \bar{P}_1(n_0) = G(u_4) - G(u_3) \tag{2.9}$$

where

$$u_1 = \frac{-n_0 E_0(z)}{\sqrt{n_o}\ \sigma_0(z)} \qquad u_3 = \frac{\log B - n_0 E_1(z)}{\sqrt{n_0}\ \sigma_1(z)}$$

$$u_2 = \frac{\log A - n_0 E_0(z)}{\sqrt{n_0}\ \sigma_0(z)} \qquad u_4 = \frac{-n_0 E_1(z)}{\sqrt{n_0}\ {}_1(z)} \tag{2.10}$$

$G(u)$ = probability that a $N(0,1)$ variable takes on a value $< u$.

$\sigma_i(z)$ = standard deviation of z when H_i is true $i = 0,1$

$E_0(z) = -\frac{1}{2}(\theta_0 - \theta_1)^2 \qquad E_1(z) = \frac{1}{2}(\theta_0 - \theta_1)^2$

The value of n_0 must be sufficiently large to insure $z_1 + \ldots + Z_{n_0} \sim N$ with mean equal to $n_0 E_i(z)$ and standard deviation equal to $\sqrt{n_0}\ \sigma_i(z)$ when H_i is true, $i = 0,1$. The upper bounds for the α and β errors are then given by:

$$\alpha(n_0) \leq \alpha + \bar{p}_0(n_0) \tag{2.11}$$

$$\beta(n_0) \leq \beta + \bar{p}_1(n_0) \tag{2.12}$$

Modifications to SPRT

As was stated earlier, Wald's SPRT will require a smaller average number of observations than any other sequential test of one simple hypothesis against another when the true value of θ is either θ_0 or θ_1. This,

is not always the case. Read [47] and Anderson [2] have proposed modifications to the SPRT in case the true value of θ lies between θ_0 and θ_1.

Read's Partial Sequential Probability Ratio Test (PSPRT) uses the same initial boundaries, A and B, as Wald's test. Initially, a fixed number, n, of observations is taken. At the end of these n observations, the quantities:

$$B_n' = B \frac{p_0{}^n}{p_1{}^n} \tag{2.13}$$

and

$$A_n' = A \frac{p_0{}^n}{p_1{}^n} \tag{2.14}$$

where p_{in} = joint likelihood of the first n observations under h_i, i=0,1 are computed. If the inequality

$$0 < B_n' < 1 < A_n' \tag{2.15}$$

does not hold, a decision is made at this point. If $A_n' < 1$, then H_1 is accepted and if $B_n' > 1$ then H_0 is accepted. If (2.15) holds, then sampling is continued as in the SPRT and the following decision rules are used:

A. Stop sampling and decide in favor of H_0 as soon as $p_1(x_{n+1},\ldots,x_n')/p_0(x_{n+1},\ldots,x_n')$ becomes less than B_n'.

B. Stop sampling and decide in favor of H_1 as soon as $p_1(x_{n+1},\ldots,x_n')/p_0(x_{n+1},\ldots,x_n')$ becomes greater than A_n'.

C. Continue sampling as long as $B_n' < p_1(x_{n+1},\ldots,x_n')/p_0(x_{n+1},\ldots,x_n') < A_n'$.

Where $p_i(x_{n+1},\ldots,x_n')$ is the joint likelihood of observations x_{n+1} to

$x_{n'} > n$ under H_i, $i = 0,1$.

Read shows that his PSPRT will require, on the average, more observations than the SPRT when the true value of θ is θ_0 or θ_1 but may require substantially fewer observations when the true value of θ is between θ_0 and θ_1, say $\theta = \frac{1}{2}(\theta_0 + \theta_1)$. In both cases, Read's Average Sample Number (ASN) is lower than the fixed sample size. Read's procedure is most useful when it is desirable to take n initial observations for some reason such as to provide the experimenter with an estimate of the sampling costs.

Anderson's modification to the SPRT applies only to densities of the Koopman-Darmois (exponential) form and he presents specific examples for a normal distribution with unknown mean and known variance. Since many random variables follow a normal distribution, or do so approximately, this is a reasonable approach.

In the SPRT, the continuation region boundaries, A and B, can be thought of as describing two parallel lines, $y = A$ and $y = B$, where the decision is made as soon as the value of y, in this case the likelihood ratio, crosses one of the lines. Anderson proposes replacing the parallel lines with a set of converging lines $y = c_1 + d_1 n$ and $y = c_2 + d_2 n$ with truncation of the sequential procedure at some value N. To avoid intersection of the lines before $n = N$, it is necessary that $c_2 + d_2(N-1) < c_1 + d_1(N-1)$. It is also desirable that the lines converge, so it is necessary that $d_1 < 0 < d_2$.

For a normal distribution, as described here, Hoel, Part and Stone [28] have shown that the likelihood ratio can be replaced by the quantity $\sum_{i=1}^{n} x_i$ with appropriate modifications to the boundaries of the continuation region. Making this substitution, the decision process for Anderson's

SPRT becomes:

A. Stop sampling and decide in favor of H_0 as soon as $\sum_{i=1}^{n} x_i$ becomes less than $c_2 + d_2 n$.

B. Stop sampling and decide in favor of H, as soon as $\sum_{i=1}^{n} x_i$ becomes greater than $c_1 + d_1 n$.

C. Continue sampling as long as $c_2 + d_2 n < \sum_{i=1}^{n} x_i < c_1 + d_1 n$.

Anderson's calculations of the ASN for his modification is extremely complicated and involve use of the Wiener stochastic process. However, his ASN for intermediate values of θ, between θ_0 and θ_1, is lower than the ASN for Read's PSPRT. Naturally, due to its optimum property, Wald's SPRT has a lower ASN at $\theta = \theta_0$ and $\theta = \theta_1$.

Testing Other Than Simple Hypotheses

The SPRT can only be used to test a simple hypothesis against a simple alternative so there are many real life situations in which it is not applicable. For example, to test whether a new product is better than an already existing product it is designed to replace would involve testing $H_0: \theta = \theta_0$ vs. $H_1: \theta > \theta_0$ and the SPRT would not apply. In this case, it is possible to modify the alternate hypothesis so that the SPRT could be used. Assuming that the new product would not replace the old one unless there was a significant difference in some performance characteristic, the alternate hypothesis can be written as $H_1: \theta - \theta_1 = d$ where d is the required margin of difference between the new and old products. Now the alternate hypothesis is in simple form and the SPRT may be applied.

Another special case is when the decision to be made is of the form: product x is inferior to product y or there is no difference between products x and y or product x is superior to product y. This is equivalent

to testing $H_{-1}:\theta < \theta_1$ vs. $H_0:\theta_2 < \theta < \theta_3$ vs. $H_1: \theta > \theta_4$ where $\theta_1 < \theta_2 < \theta_3 < \theta_4$. In the region $\theta < \theta_1$ it is desirable to decide that product X is inferior to product y; in the region $[\theta_2, \theta_3]$ there is no difference between the products; and in the region $\theta > \theta_4$ product x is superior to product y. Sobel and Wald [50] and Armitage [6] have both described methods of solving this problem.

Sobel and Wald define their three-decision problem as deciding between hypotheses $H_{-1}:\theta < a_1$, $H_0:a_1 < \theta < a_2$ and $H_1:\theta > a_2$. This definition leads to a division of the parameter space into five regions. Around a_1 there is a region (θ_1, θ_2) where there is no strong preference between H_{-1} and H_0 but where it is strongly desired to reject H_1. Around a_2 there is a region (θ_3, θ_4) where there is no strong preference between H_0 and H_1 but where it is strongly desired to reject H_{-1}. For $\theta \leq \theta_1$, the desirable decision is to accept H_{-1}; for $\theta_2 \leq \theta \leq \theta_3$, the desirable decision is to accept H_0; and for $\theta \geq \theta_4$, the desirable decision is to accept H_1. Given this formulation, a wrong decision can be made in the following manner:

A. Acceptance of H_0 or H_1 for $\theta \leq \theta_1$.

B. Acceptance of H_1 for $\theta_1 < \theta < \theta_2$.

C. Acceptance of H_{-1} or H_1 for $\theta_2 \leq \theta \leq \theta_3$.

D. Acceptance of H_{-1} for $\theta_3 < \theta < \theta_4$.

E. Acceptance of H_{-1} or H_0 for $\theta \geq \theta_4$.

Sobel and Wald then consider the case where:

A. Probability of a wrong decision $\leq \gamma_1$ for $\theta \leq \theta_1$.

B. Probability of a wrong decision $\leq \gamma_2$ for $\theta_1 < \theta < \theta_4$.

C. Probability of a wrong decision $\leq \gamma_3$ for $\theta \geq \theta_4$.

And the special case where $\gamma_1 = \gamma_2 = \gamma_3 = \gamma$ for all values of θ.

Their procedure consists of conducting two concurrent sequential probability ratio tests, R_1 and R_2. R_1 is used to test the hypotheses $\theta = \theta_1$ vs. $\theta = \theta_2$ and R_2 tests $\theta = \theta_3$ vs. $\theta = \theta_4$. Their decision procedure consists of the following:

A. Both R_1 and R_2 are computed after each trial until

B. Either: one ratio leads to a decision to stop. Then this ratio is no longer computed but the other one is until it also leads to a decision to stop.

C. Or: both ratios lead to a decision to stop at the same stage.

The final decision, R, can be made from the results shown in Table 8. Sobel and Wald give a proof that the case where R_1 accepts θ_1 and R_2 *accepts* θ_4 *can never occur.* In order to properly define the SPRT's R_1 and R_2, it is necessary to either be given values A, B, $\hat{A}$, and $\hat{B}$ which form the boundaries of the critical regions or to approximate them based on the upper bounds for the respective probabilities of making a wrong decision. These approximations are shown to be:

$$A = (1 - \gamma_2)/\gamma_1 \tag{2.16}$$

$$B = \gamma_2/(1 - \gamma_1) \tag{2.17}$$

$$\hat{A} = (1 - \gamma_3)/\gamma_2 \tag{2.18}$$

$$\hat{B} = \gamma_3/(1 - \gamma_2) \tag{2.19}$$

The special case $\gamma_1 = \gamma_2 = \gamma_3 = \gamma$ is easily handled by substituting in the above approximations.

Armitage's method is similar except that it involves using three

Table 8. Decision Process for Testing Multiple Hypotheses

	R_1		R_2		R
If	Accepts θ_1	and	Accepts θ_3	then	Accepts H_{-1}
If	Accepts θ_2	and	Accepts θ_3	then	Accepts H_0
If	Accepts θ_2	and	Accepts θ_4	then	Accepts H.

SPRT's concurrently so that all possible combinations of the three alternative hypotheses taken two at a time are tested. Armitage's decision procedure is as follows:

A. All three SPRT's are examined after each trial.

B. Sampling is continued until the results indicate that one hypothesis is preferred over both of the other two hypotheses based on the SPRT's involving those hypotheses and the preferred one.

C. The decision is then made in favor of the preferred hypothesis.

Armitage's procedure has one decided advantage over that of Sobel and Wald. By requiring that all three SPRT's be examined until completion, a definite decision in favor of one of the hypotheses will be reached. Using Sobel and Wald's procedure, it is possible to terminate sampling before a definite decision point has been reached. This situation is shown in Figure 2. For the path, OT, shown in the figure, test R is terminated as soon as line AB is crossed with the decision being made in favor of H_0. R_1 is not calculated again even though the path may go back across the line. Test R_2 is terminated as soon as line AC is crossed with, once again, the decision being made in favor of H_0. Based on Table 8,

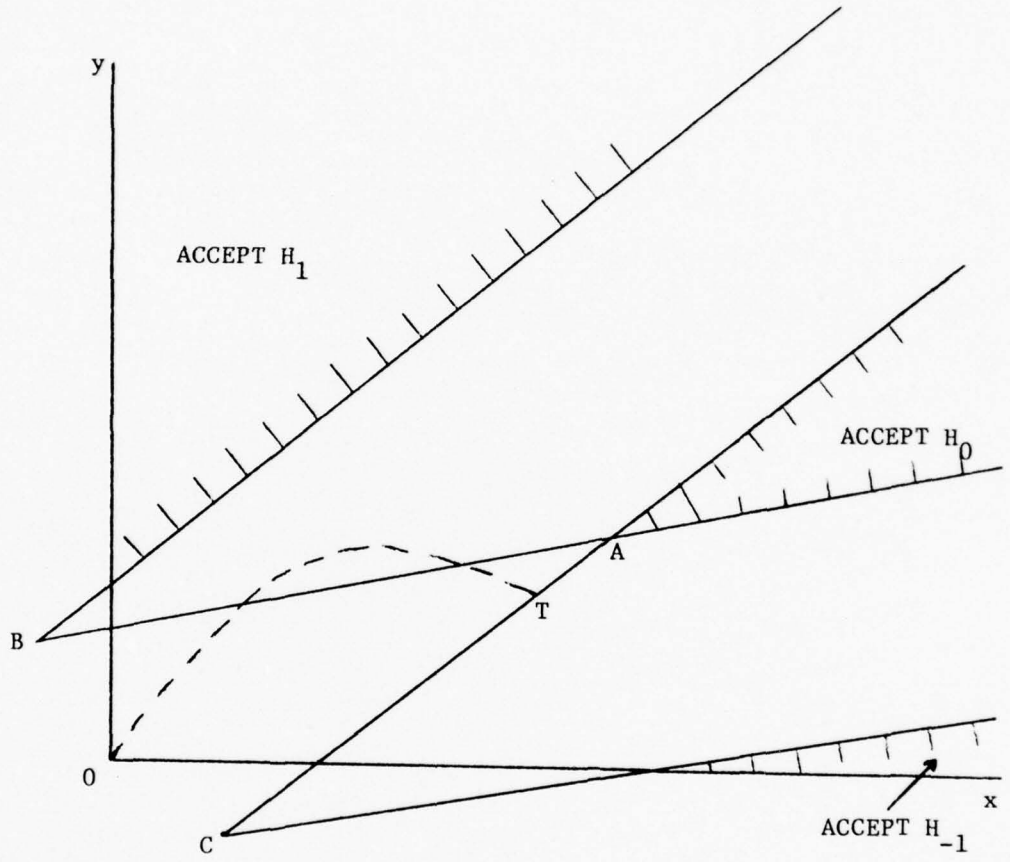


Figure 2. Testing Multiple Hypotheses.

the terminal decision would be in favor of H_0 but, as can be seen from the figure, the shaded region, H_0, was never reached.

Tests on Mean of a Normal Distribution with Unknown Variance

Sequential T-Test. Once again, the problem is to test the hypothesis $H_0:\theta = \theta_0$ vs. $H_1:\theta \geq \theta_1 > \theta_0$. Since the standard deviation, $\sigma = \sqrt{\sigma^2}$, will always be a positive number, the alternate hypothesis can be expressed as $H_1:\theta \geq \theta_1 = \theta_0 + \delta\sigma$ where δ is some positive number. This can be rewritten as $H_1:\theta - \theta_0 \geq \delta\sigma$. Since the value of σ is unknown, the problem is to find a procedure that does not depend on σ. A random variable defined as the ratio of a N(0,1) variable to the square root of a chi-square random variable divided by its degrees of freedom follows the t-distribution and the formation of such a random variable will eliminate the unknown variance. Govindarojulu [23] has formulated a sequential procedure based on the statistic:

$$t_n = (n)^{\frac{1}{2}}(\bar{x}_n - \theta_0)/S_n \tag{2.20}$$

where $\bar{x}_n$ = sample mean = $(\sum_{i-1}^{n} x_i)/n$

S_n = sample standard deviation = $(\sum_{i=1}^{n} (x_i - \bar{x})^2/(n-1))^{\frac{1}{2}}$

His procedure is as follows:

A. After each trial, compute t_n as shown in (2.20).

B. Continue sampling as long as $B_n < t_n < A_n$.

C. Stop sampling and decide in favor of H_0 if $t_n < B_n$.

D. Stop sampling and decide in favor of H_1 if $t_n > A_n$.

The boundaries A_n and B_n are arbitrary with the only requirement being that they be chosen so that the procedure terminates finitely. David

and Kruskal [17] have shown that this will occur when A_n and B_n are obtained by equating the ratio of a non-central t distribution to a central t distribution to the constants A and B from Wald's SPRT. This t-test can easily be modified to test $H_0:\theta = \theta_0$ vs. $H_1:|\theta - \theta_0| \geq \delta\sigma$ which is equivalent to $H_1:\theta \neq \theta_0$ by using the region $B_n < |t_n| < A_n$ as the region in which sampling is continued.

Two Stage T-Test. Several people have proposed modifications to the t-test that require an initial number, n_0, of observations. Govindarajulu's [36] Minimum Probability Ratio Test (MPRT) is equivalent to Anderson's modification to the SPRT only for the case where the variance is unknown. For the case where $\alpha = \beta$, Govindarajulu defines

$$C_{n_0} = (n_0 - 1)[(2\alpha)^{-2/(n_0-1)} - 1] \tag{2.21}$$

His procedure consists of the following:

A. Take n_0 initial observations and compute $S^2_{n_0}$ as defined in (2.20).

B. For each observation $n \geq n_0$, after observing the response x_n, stop sampling as soon as

$$\delta\Big|\sum_{i=1}^{n}(x_i - \delta/2)\Big|/S^2_{n_0} \geq C_{n_0} - n\delta^2/4S^2_{n_0}, \quad n \geq n_0 \tag{2.22}$$

C. Decide in favor of H_0 if $\sum_{i=1}^{n}(x_i - \delta/2) < 0$.

D. Decide in favor of H_1 of $\sum_{i=1}^{n}(x_i - \delta/2) > 0$.

E. Continue sampling if $\delta\Big|\sum_{i=1}^{n}(x_i - \delta/2)\Big|/S^2_{n_0} - n\delta^2/4S^2_{n_0}$.

In the case where $\alpha \neq \beta$, the procedure may still be used by replacing 2α with $\alpha + \beta$ in (2.21). Govindarajulu's procedure could be further

modified by computing C_{n_0} and $S^2_{n_0}$ after each observation.

Baker (9) as modified by Hall (24) has proposed another two stage procedure that will be discussed further in the next chapter.

CHAPTER III

DEVELOPMENT OF METHODOLOGY

Introduction

This chapter will present a formalized procedure for building a factorial experiment with the goal of reducing the required sample size without reducing the amount of information obtained from the experimentation. This information may include data about the significance of the various factors, the levels of significance, or the probabilities of making a type I or type II error. The method developed by Russ involves minimizing the expected additional system cost (EASC) given by

$$EASC = C_0 + \sum_{i=1}^{N} C_i + \alpha C_\alpha + \beta C_\beta \qquad (3.1)$$

where C_0 = fixed cost of testing. This can be considered a fixed cost independent of the design structure.

N = total number of observations taken.

C_i = sampling cost of i^{th} observation. This cost will be assumed to be a constant, C_s, for each sample taken so the term $\sum_{i=1}^{n} C_i$ can be rewritten as NC_s.

C_α = penalty cost for a Type I error. Assumed fixed for a given Operational Test.

α = probability of a Type I error. Set at an acceptable level by the test designer.

C_β = penalty cost for a Type II error. Assumed fixed for a given

Operational Test.

β = probability of a Type II error. Set at an acceptable level by the test designer.

For a given operational test all of the costs in the EASC equation can be considered as fixed so the equation can be minimized simply by minimizing the number of observations required to obtain the desired information.

Basic Assumptions

Since Operational Testing is designed to test one proposed system against another or against some standard for comparison, the hypothesis to be tested is of the form:

$$H_0: \mu = \mu_0 \qquad (3.2)$$
$$H_1: \mu > \mu_0$$

where μ = standard for comparison.

μ_0 = population mean for the item being tested.

Operational testing will require that the prototype item exceed the standard for comparison by a certain margin before the decision to accept the item is made so (3.2) can be rewritten as:

$$H_0: \mu - \mu_0 = 0$$
$$H_1: \mu - \mu_0 = d \qquad (3.3)$$

where d is the required margin of difference.

Since d will be some positive number, it can be expressed as some constant multiple of the population standard deviation, also a positive number, and (3.3) can be rewritten as:

$$H_0: \mu - \mu_0 = 0$$
$$H_1: \mu - \mu_0 = \delta\sigma \qquad (3.4)$$

Given the nature of operational testing, the assumption is also made that the observed responses will come from a normal population with unknown mean μ and variance σ^2. The value of σ^2 may or may not be known and each case will be treated separately. If σ^2 is known, the problem is greatly simplified. If it is not known, a sequential procedure can be developed but an alternate solution is to obtain some estimate of σ^2 as soon as possible. This estimate may be obtained from any prior testing done on the system, from comparison with a similar type system, or from the results obtained in OT I and then used throughout the remaining operational tests.

Variance Known

If the variance is known, the sequential probability ratio test can be employed to make a decision to either accept the null or alternate hypothesis or to continue sampling. Since (3.4) is in the form of a simple null hypothesis against a simple alternate hypothesis, the likelihood ratio test:

$$\frac{L_{in}}{L_{on}} = \frac{\sum_{i=1}^{n} F(x_i|\theta_1)}{\prod_{i=1}^{n} F(x_i|\theta_0)} \qquad (3.5)$$

can be used as a basis for making this decision. The sequential procedure should include a region for accepting H_0, a region for accepting H_1, and a region in which the decision to continue sampling is made. For a normal population:

$$L = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left(-\tfrac{1}{2}\left(\frac{x_i-\mu}{\sigma}\right)^2\right) = \frac{1}{(2\pi\sigma^2)}\, n/2 \quad \exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^{n} (x_i - u)^2\right) \tag{3.6}$$

and therefore

$$\frac{L_{in}}{L_{on}} = \frac{\frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^{n} (x_i-\mu_1)^2\right)}{\frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^{n} (x_i-\mu_0)^2\right)} \tag{3.7}$$

which reduces to

$$\frac{L_{in}}{L_{on}} = \exp\left(\frac{\mu_1-\mu_0}{\sigma^2} \sum_{i=1}^{n} x_i + \frac{n(\mu_0^2-\mu_1^2)}{2\sigma^2}\right) \tag{3.8}$$

The decision rule for the sequential procedure will then reduce to:

1) Stop sampling and accept H_0 if (3.8) < B
2) Stop sampling and accept H_1 if (3.8) > A (3.9)
3) Continue sampling if B < (3.8) < A

As stated in Chapter II, values of

$$B = \frac{\beta}{1-\alpha} \text{ and } A = \frac{1-\beta}{\alpha} \tag{3.10}$$

will result in a sequential test with the desired probabilities of error, α and β.

Substituting (3.10) into (3.9), taking logarithms, and simplifying yields the following sequential procedure for a normal population with known variance:

1) Stop sampling and accept H_0 if $\frac{\sigma^2}{\mu_1-\mu_0} \log \frac{\beta}{1-\alpha} + \frac{n(\mu_1+\mu_0)}{2} > \Sigma x_i$

2) Stop sampling and accept H_1 if $\frac{\sigma^2}{\mu_1-\mu_0} \log \frac{1-\beta}{\alpha} + \frac{n(\mu_1+\mu_0)}{2} < \Sigma x_i$

3) Continue sampling if

$$\frac{\sigma^2}{\mu_1-\mu_0} \log \frac{\beta}{1-\alpha} + \frac{n(\mu_1+\mu d}{2} < \Sigma\, x_i < \frac{\sigma^2}{\mu_1-\mu_0} \log \frac{1-\beta}{\alpha} + n(\frac{\mu_1+\mu_0}{2})$$

where μ_0 is as hypothesized in (3.4) and $\mu_1 = \mu_0 + \delta\sigma$

Variance Unknown

If the variance of the normal population is unknown, the problem is considerably more difficult since the likelihood ratio (3.8) will depend on the value of the unknown variance. If a reasonable approximation to the variance can be obtained, a sequential probability ratio test can be performed that will closely approximate the test for the variance known case. In most cases, this approximation will not be available. Use can then be made of the fact that a random variable formed as the ratio of a N(0,1) variable to the square root of a chi-square variable divided by its degrees of freedom follows a t distribution.

Under the assumption that the response variable, X_i, comes from a $N(\mu,\sigma^2)$ population, the sum of n observed responses will follow a $N(\mu,n\sigma^2)$ distribution and the statistic

$$Z = (\bar{X} - \mu)/\sigma/\sqrt{n} \qquad (3.11)$$

will follow a N(0,1) distribution. A chi-square random variable with n degrees of freedom is the sum of the squares of n independent N(0,1) variables so the quantity

$$\sum_{i=1}^{n} ((X_i - \mu)/\sigma)^2 \tag{3.12}$$

will be distributed as chi-square with n degrees of freedom and the t-statistic proposed by Govindarajulu in (2.20) can be formed.

The procedure of Baker [9] as modified by Hall [24] incorporates the required margin of difference, δ, in the formulation of the test statistic

$$r_n(S^2_{n_0}) = (\delta \sum_{i=1}^{n} (x_i - \delta/2))/S^2_{n_0} \qquad (n \geq n_0) \tag{3.13}$$

in the proposed two stage procedure. Baker develops the following as appropriate upper and lower bounds to the sampling region:

$$a_{n_0} = \tfrac{1}{2}(n_0-1)(\alpha^{-2/n_0-1} - 1) \tag{3.14}$$

$$b_{n_0} = -\tfrac{1}{2}(n_0-1)(\beta^{-2/n_0-1} - 1) \tag{3.15}$$

The sequential procedure for the variance unknown case then reduces to:

A. Observe the first n_0 observations of the response variable, $x_1, x_2, \ldots, x_{n_0}$.

B. Compute the quantities:

$$\bar{x}_{n_0} = \sum_{i=1}^{n_0} x_i/n_0 \tag{3.16}$$

and

$$S^2_{n_0} = \sum_{i=1}^{n_0} (x_i - \bar{x}_{n_0})^2/(n_0-1) \tag{3.17}$$

C. For each observation $n \geq n_0$, after observing X_n:

1. Stop sampling and decide in favor of H_0 if $r_n(S^2_{n_0}) < b_{n_0}$.
2. Stop sampling and decide in favor of H_1 if $r_n(S^2_{n_0}) > a_{n_0}$.
3. Continue sampling if $b_{n_0} < r_n(S^2_{n_0}) < a_{n_0}$.

The modification of recomputing $S^2_{n_0}$, a_{n_0} and b_{n_0} after each observation may be easily incorporated into the procedure. This test would then be equivalent to the SPRT (σ^2 known) with wider boundaries to account for the fact that σ^2 is unknown but the boundaries would converge to the SPRT boundaries as n becomes large.

Solution Procedure

The proposed solution algorithm combines the sequential procedures just developed with a plan for systematically building a factorial experiment in order to obtain as much information as possible from the experiment in as few observations as possible. Use is made of screening experiments in building the full factorial in order to eliminate unnecessary control variables from future experimentation.

The first step, once the requirement to conduct operational testing is received, is to determine which variables can be expected to affect the response to be measured and how much they can be controlled. This will set the size of the full experiment. Once the number of observations that can be developed by assuming that higher order interactions are negligible in order to fractionate the full factorial. Once the first fraction is run, a sequential analysis is performed to determine if further experimentation is necessary. If it is, the second fraction, with all signs reversed, is run and the sequential procedure is repeated. In

addition, analysis of the effects due to the various factors is performed to determine probable significance of effects. From this point on, each successive fraction is examined based on changing signs for appropriate factors in order to isolate those factors of interest. In case there are no factors or interactions of special interest, a heuristic approach would be to change the signs for the major factors proven to be significant to obtain as much information about them and their two factor interactions. Experimentation is terminated as soon as one of the sequential boundaries is crossed and a decision is made to accept one or another of the hypotheses. At this time, an analysis of variance can be performed to determine which effects are or are not significant.

The formalized algorithm is as follows:

1) Determine the variables of interest and the levels each are to be examined at.

2) Determine the number of observations that can be made under homogeneous conditions.

3) Determine the generating relationship to fractionate the full factorial experiment so that the number of observations in a block, n_0, is less than or equal to the number determined in 2 above.

4) Determine required input data and parameters:

   a) Acceptable levels of type I and type II errors, α and β.

   b) Required difference margin, δ.

   c) Actual value of an estimate for variance of the response variable if possible.

   d) Hypotheses to be tested.

5) Pick one of the blocks at random and perform the experimentation.

6) If the value of the variance is unknown or if there is no reasonable estimate for the variance, go to 8. Otherwise continue.

7) Perform sequential analysis using the sequential probability ratio test since the variance is known. If the SPRT results in a decision to stop sampling, go to 14; otherwise, go to 9.

8) Perform sequential analysis using the sequential t test since the variance is unknown. If the sequential t test results in a decision to stop sampling, go to 14; otherwise, continue.

9) Perform the next block of experiments with all signs in the first block reversed. If the variance is known, go to 10. If the variance is not known, go to 11.

10) Perform sequential analysis using the SPRT. If the decision is to stop sampling, go to 14. If not, go to 12.

11) Perform sequential analysis using the sequential t test. If the decision is to stop sampling, go to 14. If not, continue.

12) Perform a screening analysis of the effects due to the main factors in order to determine relative significance of main factors and combinations of two-factor interactions.

13) Determine factors and interactions of interest and perform the next block of experiments found by switching the signs in the column for the factor of interest. If the variance is known, go to 10. If the variance is not known, go to 11.

14) Perform an ANOVA to determine the level of significance of all factors of interest. Eliminate non-significant factors from further testing by setting them at some common level.

15) If sequential analysis does not result in a decision in favor of one

hypothesis or another, stop sampling and perform an ANOVA after one observation has been made at each treatment combination.

The proposed methodology will be illustrated in the next chapter by means of an example based on an actual operational test and the sensitivity of the input parameters demonstrated.

CHAPTER IV

DEMONSTRATION OF THE APPROACH

Introduction

This chapter will present an example of the methodology developed in the previous chapter. Hypothetical results from an Operational Test are presented and analyzed in the classical manner currently employed. Then the proposed methodology is applied to the same results with a reduction in the sample size required to gain the same information from the data. Sensitivity analysis of the input parameters and basic assumptions are also performed to demonstrate the robustness of the procedure.

Background

The Commander, U.S. Army Operational Test and Evaluation Agency (OTEA) has been given the requirement to conduct Operational testing to evaluate the overall effectiveness of the new Artillery Locating Radar, AN-TPQ-37, designed as a replacement for the system currently in use.

The test plan calls for testing several different performance aspects of the Artillery Locating Radar (ALR), one of which is its ability to detect hostile artillery fire. Since the radar cannot locate hostile artillery unless it first detects it, the most critical issue for this test is the percentage of hostile artillery rounds detected. The manufacturer has determined that five factors should influence the performance of the radar. They are the threat array employed by the enemy, the use of electronic counter measures (ECM) by the enemy, the rate of hostile

fire, the range from the ALR to the enemy threat, and the sector the ALR is searching. All of these factors can be set at two levels except for the threat array. That can be set at one of four levels to represent the typical composition of enemy artillery units that the ALR is likely to be deployed against. Since four is a multiple of two, the threat array factor can be represented as two pseudo-factors each with two levels and the entire experiment can be represented as a 2^6 factorial experiment. The six factors and the levels of each factor are shown in Table 9.

Table 9. Factors in 2^6 Factorial Experiments

Factor	Low Level	High Level
A - ECM	Not Employed	Employed
B - Rate of Fire	Slow	Fast
C - Range	Short $\leq$ 10,000 KM	Long $\geq$ 10,000 KM
D - Sector	Narrow ($\pm$ 15° of center	Wide (15-45° of center)
E - Threat Array	I	II
F - Threat Array	III	IV

Where THREAT ARRAY I = 1 enemy battery
II = 2 enemy batteries
III = 1 enemy battalion
IV = 2 enemy battalions

The Commander, OTEA, has stated that the ALR be tested by determining the percentage of hostile artillery round detected from firings taken at each of the 64 possible treatment combinations. The analysis

method used will be Analysis of Variance (ANOVA) with all third order and higher interactions assumed to be negligible and pooled to form an estimate of the error. The standard for comparison is the current system which is detecting 50% of all hostile artillery rounds. The measure of effectiveness employed will be the percentage of rounds located. The principle purpose of the test will be to determine whether or not the mean percentage of rounds detected by the ALR exceeds the current system by some multiple of the standard deviation. It is assumed that the percentage of rounds detected comes from a normal distribution with unknown mean, μ, and variance, σ^2. Therefore, the testing will consist of a test of $H_0: \mu = \mu_0$ vs. $H_1: \mu = \mu_0 + \delta\sigma$, where δ is some positive constant.

Classical Methods

OTEA conducted test firings at all 64 treatment combinations and determined the percentage of rounds detected from each firing. Time limitations precluded any more than eight test firings per day so a 2^{6-3} resolution III design with I = ABD = ACE = BCF = BCDE = ACDF = ABEF = DEF as the generating relationship was used to fractionate the design and then the eight different blocks were fired in random order. The results, percentage of rounds detected, and the ANOVA for the entire experiment are shown in the appendices. The results of the ANOVA indicate that only the factors relating to ECM, rate of fire, and sector of search are significant so OT II should be conducted using only those factors.

Proposed Methodology

Using the same data and restrictions as in the actual OT I, the

following example will illustrate the proposed sequential procedure.

The first eight observations are obtained from the block generated by I = ABD = ACE = BCF = BCDE = ACDF = ABEF = DEF and consists of treatment combinations def (37), af (36) be (40), abd (87), cd (61), ace (34), bcf (46) and abcdef (89) where the values in parentheses represent the observed response for that treatment combination. Based on these observed responses, $\delta = 1$, $\alpha = .10$, and $\beta = .10$, the following computations are performed:

$$\bar{y} = 430/8 = 53.75 \tag{4.1}$$

$$S_8^2 = 3635.5/7 = 519.35 \tag{4.2}$$

$$A_8 = \tfrac{1}{2}(7)(.1^{-2/7}-1) = 3.26 \tag{4.3}$$

$$B_8 = -\tfrac{1}{2}(7)(.1^{-2/7}-1) = -3.26 \tag{4.4}$$

$$r_8(S_8^2) = \delta(\Sigma(x_i - \delta/2))/S_8^2 = 426/519.35 = .82 \tag{4.5}$$

Since $B_8 < r_8(S_8^2) < A_8$, the decision is made to continue sampling.

The alias structure for this 2^{6-3} resolution III design is shown in Table 10. Assuming that all third order and higher interactions are negligible, each major factor is aliased with two two-factor interactions. In order to separate the main effects from their aliased two-factor interactions, the next block run will be the same as the first but with all signs reversed. The defining relationship for this block will then be I = -ABD = -ACE = -BCF = BCDE = ACDF = ABEF = -DEF. The block consists of the following treatment combinations and their observed responses: abc (69), bcde (56), acdf (67), cef (10), abef (69), bdf (56), ade (75),

Table 10. Alias Structure for 2^{6-3}_{III} Design

Factor	A L I A S E S						
I	ABD	ACE	BCF	BCDE	ACDF	ABEF	DEF
A	BD	CE	ABCF	ABCDE	CDF	BEF	ADEF
B	AD	ABCE	CF	CDE	ABCDF	AEF	BDEF
C	ABCD	AE	BF	BDE	ADF	ABCEF	CDEF
D	AB	ACDE	BCDF	BCE	ACF	ABDEF	EF
E	ABDE	AC	BCEF	BCD	ACDEF	ABF	DF
F	ABDF	ACEF	BC	BCDEF	ACD	ABE	DE
AF	BDF	CEF	ABC	ABCDEF	CD	BE	ADE

and (1)(24). Sequential calculations result in the following:

$$\bar{y} = 856/16 = 53.5 \tag{4.6}$$

$$S^2_{16} = 502.4 \tag{4.7}$$

$$A_{16} = 2.695 \quad , \quad B_{16} = -2.695 \tag{4.8}$$

$$r_{16}(S^2_{16}) = 848/502.4 = 1.69 \tag{4.9}$$

Since $B_{16} < r_{16}(S^2_{16}) < A_{16}$, the decision is made to continue sampling. Before the next block is run, however, an analysis of the results obtained so far is performed. The results are shown in the appendices. Based on the relative magnitude of the effects, it would appear that factors A, B, and D were significant while factors C, E. and F were not. Since the largest interaction term is due to BD and CE, it is also

possible that the BD interaction is significant. Since the decision has already been made to continue sampling, the next block run will be the same as the first block except that all of the signs in the column for factor B will be reversed. This will isolate factor B and all of its two factor interactions and will provide a better estimate of whether or not the BD interaction is significant.

The third block then consists of the following treatment combinations and their observed responses: bdef (67), abf (74), e (19), ad (66), bcd (67), abce (72), cf (11), acdef (75). The sequential calculations result in the following:

$$\bar{y} = 1298/24 = 54.08 \tag{4.10}$$

$$S_{24}^2 = 563.99 \tag{4.11}$$

$$A_{24} = 2.55 \qquad B_{24} = -2.55 \tag{4.12}$$

$$r_{24}(S_{24}^2) = 1286/563.99 = 2.28 \tag{4.13}$$

Since $B_{24} < r_{24}(S_{24}^2) < A_{24}$, the decision is made to continue sampling. An analysis of the results of the data from the third block is shown in the appendices. Once again, factors A, B, and D appear to be significant while the BD interaction is still in doubt. After the second block, the effect in question was due to the BD and CE interactions so it is possible that this was due to the CE rather than the BD interaction. Since the decision has been made to continue sampling, the next block run will change the signs in the column for factor c to isolate factor c and its two factor interactions. This block contains the following treatment combinations and observed responses: cdef (48), acf (41), bce (30),

abcd (95), d (57), ae (46), bf (31), and abdef (86). The sequential calculations result in the following:

$$\bar{y} = 1732/32 = 54.125 \tag{4.14}$$

$$S_{32}^2 = 550.31 \tag{4.15}$$

$$A_{32} = 2.48 \quad , \quad B_{32} = -2.48 \tag{4.16}$$

$$r_{32}(S_{32}^2) = 1716/550.31 = 3.12 \tag{4.17}$$

Since $r_{32}(S_{32}^2) > A_{32}$, the decision is made to stop sampling and accept $H_1: \mu = \mu_0 + \delta\sigma = 50 + \sigma$. This means that the ALR has performed better to date than the current system. The analysis of the data from the fourth block is shown in the appendices. Once again factors A, B, and D appear to be significant while the CE interaction does not. Since the decision has been made to stop sampling, the recommendation made at this point would be to perform OT II using only the factors for ECM, rate of fire, and sector. The other factors would be set at some acceptable standard level and left there.

An interesting sidelight illustrated by this example is the process of collapsing a design in k variables to a smaller design in $p < k$ variables by eliminating non-significant variables from consideration. To illustrate this procedure, the experimenter was relatively sure that the variables C, E, and F were not significant after the second block of eight experiments had been performed and the screening analysis conducted. Had the decision been made to eliminate those variables from further consideration at that point by setting them at some

standard level, the next blocks of eight experiments performed would consist of a replicated 1/2 fraction of a 2^3 factorial in the significant variables A, B, and D. This is shown in the table of plus and minus signs in Table 11. In addition, the two blocks combined form a replicated full factorial in the three factors A, B, and D.

Table 11. Collapsed Design in Three Variables

BLOCK 3				BLOCK 4			
Treatment Combination	A	B	D	Treatment Combination	A	B	D
bdef	-	+	+	cdef	-	-	+
abf	+	+	_	acf	+	-	-
e	-	-	-	bce	-	+	-
ad	+	-	+	abcd	+	+	+
bcd	-	+	+	d	-	-	+
abce	+	+	-	ae	+	-	-
cf	-	-	-	bf	-	+	-
acdef	+	-	+	abdef	+	+	+

Sensitivity

This section will demonstrate how the conclusions reached in the sequential analysis would vary should any of the input parameters be changed.

α and β

Table 12 shows the different decisions made if the probabilities of Type I (α) and Type II (β) errors are changed. These results would appear to be intuitively correct. As the values for α and β are increased

the boundaries for making a decision come closer together since the experimenter is more willing to make an error while for smaller values of α and β, the boundaries are farther apart because the experimenter is willing to make an error.

Table 12. Sensitivity of α and β Errors

α,β	n = no. of Observations	A_N, B_N		$r_n(S_n^2)$	Decision
.05	8	4.74	-4.74	.82	Continue
	16	3.68	-3.68	1.69	Continue
	24	3.42	-3.42	2.28	Continue
	32	3.30	-3.30	3.12	Continue
.15	8	2.52	-2.52	.82	Continue
	16	2.16	-2.16	1.69	Continue
	24	2.06	-2.06	2.28	Stop-Accept H_1
	32	2.02	-2.02	3.12	N/A
.20	8	2.04	-2.04	.82	Continue
	16	1.80	-1.80	1.69	Continue
	24	1.73	-1.73	2.28	Stop-Accept H_1
	32	1.70	-1.70	3.12	N/A

Sequential Parameters

The sequential procedure employed involves recomputing A_{n_0}, B_{n_o}, and $S^2_{n_0}$ after each block of observations. Table 13 shows the change in the results if this modification is not employed by using the values $A_8 = 3.26$, $B_8 = -3.26$ and $S_8^2 = 519.35$ throughout the sequential analysis

Table 13. Sensitivity of Input Parameters

n	S_n^2	$r_n(S_8)$	Decision	$r_n(S_n)$	Decision
8	519.35	.82	Continue	.82	Continue
16	502.4	1.63	Continue	1.69	Continue
24	563.99	2.48	Continue	2.28	Continue
32	550.31	3.30	Stop-Accept H_1	3.12	Continue

or is only partially employed by using the values for A_8 and B_8 throughout but recomputing S_n^2 after each block. Eliminating the convergence of the boundaries as more samples are taken results in the decision to continue sampling after 32 samples have been taken but would result in the decision to stop sampling and accept H_1 after the next block of eight for a total of 40 samples.

Improvement Required (δ)

This example was run with the new system required to outperform the old system by a factor of one standard deviation. This value was arbitrarily selected by the test designer. Table 14 shows the changes in the decision making procedure as a result of changing the value of δ.

Variance Known

The preceding example illustrated the sequential analysis procedure when the observations were assumed to have come from a normal population with unknown mean and variance. In many real life cases, there may be some prior data available so that a good estimate for the variance may be obtained. Assuming that the previous testing had been conducted, the

Table 14. Sensitivity of δ

n	δ	$r_n(S_n^2)$	Decision
8	.5	.41	Continue
	1.5	1.22	Continue
	2.0	1.63	Continue
16	.5	.85	Continue
	1.5	2.52	Continue
	2.0	3.34	Stop-Accept H_1
24	.5	1.14	Continue
	1.5	3.40	Stop-Accept H_1
32	.5	1.57	Continue

MS_{Error} = 52.35 could be used as an estimate of the variance for future testing. The problem then resolves to testing $H_0:\mu = 50$ vs. $H_1:\mu = 50 + 1\sigma \simeq 57.2$ and the Sequential Probability Ratio Test can be employed. Using $\alpha = \beta = .10$, the boundaries for the SPRT become:

$$B = \frac{\beta}{1-\alpha} = \frac{.1}{.9} = .11 \tag{4.18}$$

$$A = \frac{1-\beta}{\alpha} = \frac{.9}{.1} = 9 \tag{4.19}$$

As shown in Chapter III, the SPRT consists of the following:

1) Continue sampling as long as

$$\frac{\sigma^2}{\mu_1-\mu_0}\log B + \frac{n(\mu_1+\mu_0)}{2} \leq \sum y_i \leq \frac{\sigma^2}{\mu_1-\mu_0}\log A + \frac{n(\mu_1+\mu_0)}{2}$$

which reduces to

$$\frac{52.35}{7.2}\log .11 + \frac{n(107.2)}{2} \leq \sum y_i \leq \frac{52.35}{7.2}\log 9 + \frac{n(107.2)}{2} \quad \text{or}$$

$$-16.05 + 53.6n \leq \sum y_i \leq 15.98 + 53.6n.$$

2) Stop sampling and accept H_0 if $\sum y_i < -16.05 + 53.6n$.

3) Stop sampling and accept H_1 if $\sum y_i > 15.98 + 53.6n$.

Applying the SPRT to the data from the previous example results in the following:

1) After eight observations:

$$\sum y_i = 430 \tag{4.20}$$

$$B_8 = -16.05 + (53.6)(8) = 412.75 \tag{4.21}$$

$$A_8 = 15.98 + (53.6)(8) = 444.78 \tag{4.22}$$

Since $B_8 < \sum y_i < A_8$, the decision to continue sampling is made.

2) After sixteen observations:

$$\sum y_i = 856 \tag{4.23}$$

$$B_{16} = -16.05 + (53.6)(16) = 841.55 \tag{4.24}$$

$$A_{16} = 873.58 \tag{4.25}$$

Since $B_{16} < \sum y_i < A_{16}$ the decision to continue sampling is made.

3) After 24 observations:

$$\Sigma\, y_i = 1298 \tag{4.26}$$

$$B_{24} = (-16.05) + (53.6)(24) = 1270.35 \tag{4.27}$$

$$A_{24} = 15.98 + (53.6)(24) = 1302.38 \tag{4.28}$$

Since $B_{24} < \Sigma\, y_i < A_{24}$, the decision to continue sampling is made.

4) After 32 observations:

$$\Sigma\, y_i = 1732 \tag{4.29}$$

$$B_{32} = -16.05 + (53.6)(32) = 1699.15 \tag{4.30}$$

$$B_{32} = 15.98 + (53.6)(32) = 1731.18 \tag{4.31}$$

Since $\Sigma\, y_i > A_{32}$, the decision to stop sampling and accept H_1 is made. The Sequential Probability Ratio Test resulted in the same decision after the same number of observations as the sequential t-test. An explanation for the fact that the test did not terminate sooner when the variance was assumed known can be found in the fact that the mean of the 64 observations used was 53.25 which is close to half-way between the two hypothetical values of 50 and 57.2 and the SPRT performs best at values near the hypothesized values and worse at values close to the mid-point of the hypothesized values. In spite of this fact, the SPRT still terminated in half the number of observations required by the classical methods currently employed.

CHAPTER V

RELATED APPLICATIONS

Introduction

The proposed methodology developed in Chapter III was applied to a specific situation in Chapter IV. The situation described in Chapter IV, although specific in nature, was really taken from a general class of problems, 2^n factorials, to which the methodology may be applied. This chapter will present several highly specific procedures which may be combined with the proposed methodology in certain situations to gain even more benefits from the procedure.

Major and Minor Variables

In many situations, prior knowledge of the system or of a similar system may allow the experimenter to determine which factors will definitely make a significant contribution prior to the start of experimentation. There may be several other factors about whose contribution the experimenter is unsure and, therefore desires further information. By classifying the a priori significant variables as major variables and the remainder as minor variables, use can be made of the properties of blocking a factorial experiment to gain more information from the experiment. Generally, the experimenter will desire an estimate of the main effect and all interactions for the major variables but he will be willing to assume that all interactions involving minor variables are negligible so will only want an estimate of the main effect for the minor

variable. The procedure involving major and minor variables is specific in that it requires a resolution IV design blocked into a number of blocks equal to the number of major variables. By associating each major variable with a block and the minor variables with the experimental variables, significance of a block effect will indicate significance of a major variable and an effect due to an interaction between blocks will indicate an interaction between major variables. It may be that an estimate for an effect will involve a combination of a major variable and some number of minor variable interactions but this situation poses no problem since the experimenter has assumed that all order interactions involving minor variables are negligible. A specific example of the use of major and minor variables is the 2^{8-4} resolution IV design which can be used to investigate eight minor and three major variables. Since the design contains 16 points, the first step in constructing the design matrix is to write down a full 2^4 factorial design in four of the minor variables. The remaining four minor variables are then expressed as three factor interactions of the first four variables and the major or blocking variables are expressed as two factor interactions of some pair of minor variables. The design matrix for this design is shown in Table 15. The design is then separated into eight blocks of two runs each by combining the treatment combinations that have the same signs on B_1, B_2, and B_3, i.e., the sets (-,-,-), (+,-,-), (-,+,-), (+,+,-), (-,-,+), (+,-,+), (-,+,+), and (+,+,+), to form the eight blocks. One interesting thing to note when the treatment combinations are paired in this manner is that the two treatment combinations in the same block have opposite signs for every factor.

Table 15. Design Matrix for Major and Minor Variables

Minor Variables								Major Variables		
A	B	C	D	E=ABC	F=ABD	G=ACD	H=BCD	B_1=AB	B_2=AC	B_3=AD
-	-	-	-	-	-	-	-	+	+	+
+	-	-	-	+	+	+	-	-	-	-
-	+	-	-	+	+	-	+	-	+	+
+	+	-	-	-	-	+	+	+	-	-
-	-	+	-	+	-	+	+	+	-	+
+	-	+	-	-	+	-	+	-	+	-
-	+	+	-	-	+	+	-	-	-	+
+	+	+	-	+	-	-	-	+	+	-
-	-	-	+	-	+	+	+	+	+	-
+	-	-	+	+	-	-	+	-	-	+
-	+	-	+	+	-	+	-	-	+	-
+	+	-	+	-	+	-	-	+	-	+
-	-	+	+	+	+	-	-	+	-	-
+	-	+	+	-	-	+	-	-	+	+
-	+	+	+	-	-	-	+	-	-	-
+	+	+	+	+	+	+	+	+	+	+

Sequential Factorial Estimation

In certain cases, the experimenter may desire further information about the model that represents the system under investigation. For an experiment involving P variables and assuming that all third order and higher interactions are negligible, the general model can be written as:

$$E(y) = \beta_0 + \sum_{i=1}^{P} \beta_i x_i + \sum_{i>} \sum_{j=1}^{P} \beta_{ij} x_i x_j \qquad (5.1)$$

Least squares estimates of all of the coefficients, β's, in the model can be obtained from

$$\hat{B} = (x'x)^{-1}x'y \qquad (5.2)$$

where $\hat{B} = \begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \cdots \\ \hat{\beta}_p \\ \hat{\beta}_{12} \\ \cdots \\ \hat{\beta}_{p-1}\hat{\beta}_p \end{bmatrix}$ $X = \begin{bmatrix} 1 & D \end{bmatrix}$

D = design matrix of experiment and y = column vector of observations as long as the number of observations is greater than or equal to the number of coefficients for which an estimate is desired. On the high speed computers available today, equation (5.2) can be evaluated quickly and easily.

Hunter [31] has developed a similar method where the computations may be made easily on a hand calculator in case the experimenter does not have ready access to a computer. His method requires that:

1) The model contain no more than $q \leq N$ coefficients and an experimental design containing N experiments has been completed.
2) The estimates provided by prior blocks must be mutually orthogonal with variance equal to $\frac{\sigma_2}{mN}$.

3) The added row vectors must be row-wise orthogonal ($r_i^T r_j = 0$ for $i \neq j$). Conditions 2 and 3 above are satisfied by a 2^{k-p} factorial design if $q = N$. Once an initial block of N runs has been obtained, initial estimates of the coefficients can be obtained from equation (5.2). From then on, Hunter's Predictor-Corrector equation can be used to update the estimates of the coefficients after each run. The P-C equation is given by:

$$B^* = B + \frac{1}{mN+q} \sum_{i=1}^{m} (y_i - \hat{y}_i) r_i \qquad (5.3)$$

where: B^* = $(q \times 1)$ vector of revised estimates.

B = $(q < 1)$ vector of estimates provided by prior block(s).

N = number of runs in a block.

n = number of runs completed in current block.

q = number of coefficients in the model.

m = number of blocks of N runs completed.

r_i^T = $(1 \times q)$ row vector in matrix of independent variables associated with i^{th} experiment $i = 1,2,\ldots,n \leq N$.

y_i = new observation associated with r_i^T.

$\hat{y}_i = r_i^T B$ = predicted response for i^{th} experiment.

Once the initial estimates, B, have been obtained, they are used throughout the next block. The revised estimates B^* are computed after each experimental run using the B computed after the previous block. The estimates, B, are updated after each complete block is finished and the updated estimates are then used throughout the entire next block. The variance of each revised estimate can be obtained from:

$$V(b^*) = \frac{1}{mN} \left[1 - \frac{n}{mN+q}\right] \sigma^2 \qquad (5.4)$$

where σ^2 is the population variance and m, N, n, q are as defined previously. The procedure requires computing the inverse of a matrix only after the initial block of N observations have been made.

This procedure is particularly useful when, for some reason, it becomes impossible to complete the experimentation. The change (increase) in the sum of squares due to error (deviation) for each experimental run can be computed from

$$\Delta SSD = \frac{mN}{mN+q} (y_i - \hat{y}_i)^2 \tag{5.5}$$

so that the analysis of variance table can be updated at the completion of each run. This provides the experimenter with a valid ANOVA table in the event that a complete block of N experiments cannot be completed. The complete ANOVA after m blocks of N experiments is:

Source	SS	DF
SSY = Crude SS	$\sum_{i=1}^{mN} y_i^2$	mN
SSR = Regression SS	$\sum_{i=1}^{mN} \hat{y}_i^2$	q
SSD = Error SS	$\sum_{i=1}^{mN} (y_i - \hat{y}_i)^2$	mN-q

(5.6)

To satisfy the requirement that q = N it may be necessary to introduce some slack variables. The easiest way to do this is to pick some higher order interaction(s) that may be of interest and include them in the regression model so that an estimate of their coefficients will also be obtained. An example applying this procedure to the data

from the example in Chapter IV is shown in the appendices.

Blocking Fractional Factorials

In the example illustrated in Chapter IV, each set of eight experiments consisted of a 1/8 fraction of the full 2^6 factorial. After two sets of eight experiments had been performed, the experimenter had, in effect, performed a 1/4 fraction of the 2^6 in two blocks and the experiment could be analyzed as such at this point.

To illustrate this procedure, consider the first two sets of eight experiments performed in the example in Chapter IV. They now form a 1/4 fraction of the 2^6 or a 2^{6-2} with generating relation I = BCDE = ACDF = ABEF. The alias structure for this design is shown in Table 16. Since there are two alias sets containing only three factor interactions, one of these is confounded with the two blocks run. In this case, the interaction ABD and its aliases were confounded with blocks. The eight treatment combinations in the first block performed all contain an odd number of letters in common with ABD and the eight treatment combinations in the second block performed all contain an even number of letters in common with ABD. The experimenter, in performing an Analysis of Variance at this point could then extract one degree of freedom for blocks. An analysis of the effect due to blocking could provide the experimenter with some idea of a training or learning process or an effect such as weather that may be having an effect from block to block.

Table 16. Alias Structure for 2^{6-2} with I = BCDE = ACDF = ABEF

I	BCDE	ACDF	ABEF
A	ABCDE	CDF	BEF
B	CDE	ABCDF	AEF
C	BDE	ADF	ABCEF
D	BCE	ACF	ABDEF
E	BCD	ACDEF	ABF
F	BCDEF	ACD	ABE
AB	ACDE	BCDF	EF
AC	ABDE	DF	BCEF
AD	ABCE	CF	BDEF
AE	ABCD	CDEF	BF
AF	ABCDEF	CD	BE
BC	DE	ABDF	ACEF
BD	CE	ABCF	ADEF
ABC	ADE	BDF	CEF
ABD	ACE	BCF	DEF

CHAPTER VI

CONCLUSIONS AND RECOMMENDATIONS

Limitations of the Research

This research is limited in application to univariate response models. It is assumed that the response comes from a normal population with unknown mean, μ, and variance, σ^2, which may or may not be known. The approach is demonstrated only for a 2^n factorial experiment that is fractionated into a Resolution III design but is easily extended into a factorial experiment where the factors take on any number of levels as long as the full factorial can be fractionated into a Resolution III or higher design.

Conclusions

This research accomplished three objectives:

A. An approach to systematically building a factorial experiment through the use of screening experiments was demonstrated. This allows the experimenter to obtain as much information as possible from a fixed set of resources.

B. A method of sequentially analyzing the data from a fractionated factorial experiment was demonstrated. This allows the experimenter to obtain a fixed amount of information from a reduced set of resources.

C. The proposed methodology combined the above two methods to systematically build a factorial experiment while conducting a sequential analysis of the data at the end of each block of the factorial experiment. This allows the experimenter to gain the maximum amount

of information from a minimum amount of resources.

Recommendations

Although this research considers only a univariate response model, operational testing often involves the testing of several Measures of Effectiveness (MOE). Therefore, it is recommended that future research in this area be directed at the development of a methodology to handle the case of multiple response models.

This research also was demonstrated only for a 2^n factorial experiment. Future work in the area should be directed at applying the methodology to experiments in which the factors appear at other than two levels or to other than factorial experimental designs.

As the results of more operational tests become available, it is recommended that the U.S. Army Operational Test and Evaluation Agency apply the proposed methodology to the completed test data as a further test of its validity and worth as a viable analysis method for their eventual adoption.

APPENDICES

APPENDIX A

DATA FROM THE FULL 2^6 FACTORIAL EXPERIMENT AND ANALYSIS OF VARIANCE TABLE

This appendix contains the data from the full 2^6 factorial experiment and an Analysis of Variance Table for the entire data set.

Data from the Full 2^6 Factorial Experiment

(1)	24	e	10	f	18	ef	18
a	38	ae	46	af	36	aef	34
b	33	be	40	bf	31	bef	32
ab	56	abe	74	abf	74	abef	69
c	15	ce	13	cf	11	cef	10
ac	42	ace	34	acf	41	acef	43
bc	34	bce	30	bcf	46	bcef	39
abc	69	abce	72	abcf	47	abcef	44
d	57	de	53	df	45	def	37
ad	66	ade	75	adf	66	adef	60
bd	79	bde	70	bdf	56	bdef	67
abd	67	abde	83	abdf	88	abdef	86
cd	61	cde	49	cdf	47	cdef	48
acd	78	acde	72	acdf	67	acdef	75
bcd	67	bcde	56	bcdf	65	bcdef	70
abcd	95	abcde	89	abcdf	82	abcdef	89

Analysis of Variance Table for the Entire Data Set

ANOVA

Source	SS	DF	MS	F_0
ECM	8695.56	1	8695.56	180.5**
Rate of Fire	6201.56	1	6201.56	128.7**
Range	1.00	1	1.00	.02
Sector	14460.06	1	14460.06	300.2**
Threat Array*	286.375	3	95.46	1.98
Error	2697.445	56	48.16	
Total	23342.	63		

*$SS_{\text{Threat Array}} = SS_E + SS_F + SS_{EF}$

**Significant at 1% level.

APPENDIX B

SCREENING ANALYSIS

This appendix contains an analysis of the results after each block of eight experiments was run and a comparison of the effects to determine relative significance.

A. Block 1:

$\ell_A = A + BD + CE = 62/3 = 20.67$

$\ell_B = B + AD + CF = 31.33$

$\ell_C = C + AE + BF = 10.00$

$\ell_D = D + AB + EF = 39.33$

$\ell_E = E + AC + DF = -10.00$

$\ell_F = F + BC + DE = -4.67$

$\ell_{CD} = CD + BE + AF = 4.00$

B. Block 2:

$\ell'_A = -A + BD + CE = -44.67$

$\ell'_B = -B + AD + CF = -24.67$

$\ell'_C = -C + AE + BF = 7.33$

$\ell'_D = -D + AB + EF = -27.33$

$\ell'_E = -E + AC + DF = 2.00$

$\ell'_F = -F + BC + DE = 7.33$

$\ell'_{CD} = CD + BE + AF = 18.67$

C. Block 1:

$\frac{1}{2}(\ell_i + \ell'_i)$

$BD + CE = 12.00$

$AD + CF = 3.33$

$AE + BF = 8.67$

$AB + EF = 6.00$

$AC + DF = -4.00$

$BC + DE = 1.33$

C. Block 2

$\frac{1}{2}(\ell_i - \ell'_i)$

$A = 32.67$

$B = 28.00$

$C = 1.33$

$D = 33.33$

$E = -6.00$

$F = -6.00$

D. Block 3:

$\ell_A^* = A - BD + CE = 44.00$

$\ell_B^* = -B + AD + CF = -39.33$

$\ell_C^* = C + AE - BF = 2.67$

$\ell_D^* = D - AB + EF = 36.00$

$\ell_E^* = E + AC + DF = 2.00$

$\ell_F^* = F - BC + DE = 4.00$

$\ell_{CD}^* = CD = BE + AF = 3.33$

D. Blocks 1 and 3:

$\frac{1}{2}(\ell_i + \ell_i^*)$	$\frac{1}{2}(\ell_i - \ell_i^*)$
$A + CE = 32.34$	$BD = -11.67$
$AD + CF = -4.00$	$B = 35.33$
$C + AE = 6.34$	$BF = 3.67$
$D + EF = 37.67$	$AB = 1.67$
$F + DE = -.34$	$BC = -4.34$
$CD + AF = 3.67$	$BE = .34$

E. Block 4:

$\ell_A' = A + BD - CE = 34.00$

$\ell_B' = B + AD - CF = 16.67$

$\ell_C' = -C + AE + BF = 2.00$

$\ell_D' = D + AB + EF = 46.00$

$\ell_E' = E - AC + DF = -4.67$

$\ell_F' = F - BC + DE = -7.33$

$\ell_{CD}' = -CD + BE + AF = -2.00$

F. Block 1:

$\frac{1}{2}(\ell_i + \ell'_i)$

A + BD = 27.34

B + AD = 24.00

AE + BF = 6.00

D + AB + EF = 42.67

E + DF = -7.34

F + DE = -6.00

AF + BC = 1.00

F. Block 4:

$\frac{1}{2}(\ell_i - \ell'_i)$

CE = -6.67

CF = 7.33

C = 4.00

H. O. TERMS = -3.33

AC = -2.67

BC = 1.33

CD = 3.00

APPENDIX C

SEQUENTIAL FACTORIAL ESTIMATION

This appendix applies the sequential factorial estimation procedure discussed in Chapter V to the data from the example used in Chapter IV in order to demonstrate its use.

A. After the first block of eight observations is made, estimates for the coefficients in the model:

$$y = b_0 + b_1A + b_2B + b_3C + b_4D + b_5E + b_6F + b_{24}BD \qquad (C.1)$$

can be obtained from the equation

$$\hat{\beta} = (x'x)^{-1}x'y \qquad (C.2)$$

For the first block of eight observations,

$$X = \begin{bmatrix} 1 & -1 & -1 & -1 & 1 & 1 & 1 & -1 \\ 1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 \\ 1 & -1 & 1 & -1 & -1 & 1 & -1 & -1 \\ 1 & 1 & 1 & -1 & 1 & -1 & -1 & 1 \\ 1 & -1 & -1 & 1 & 1 & -1 & -1 & -1 \\ 1 & 1 & -1 & 1 & -1 & 1 & -1 & 1 \\ 1 & -1 & 1 & 1 & -1 & -1 & 1 & -1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}, \quad y = \begin{bmatrix} 37 \\ 36 \\ 40 \\ 87 \\ 61 \\ 34 \\ 46 \\ 89 \end{bmatrix} \qquad (C.3)$$

Resulting in

$$\hat{\beta} = \begin{bmatrix} 53.75 \\ 7.75 \\ 11.75 \\ 3.75 \\ 14.75 \\ -3.75 \\ -1.75 \\ 7.75 \end{bmatrix} \qquad \text{(C.4)}$$

so the fitted model is:

$$y = 53.75 + 7.75A + 11.75B + 3.75C + 14.75D - 3.75E - 1.75F + 7.75B \qquad \text{(C.5)}$$

Since the variance-covariance matrix, $(x'x)^{-1}$ is

$$(x'x)^{-1} = \begin{bmatrix} 1/8 & & & & & & & \\ & 1/8 & & & & & & \\ & & 1/8 & & & & 0 & \\ & & & 1/8 & & & & \\ & & & & 1/8 & & & \\ & & & & & 1/8 & & \\ & & 0 & & & & 1/8 & \\ & & & & & & & 1/8 \end{bmatrix} \qquad \text{(C.6)}$$

all of the coefficients in (C.5) have variance equal to $\sigma^2/8$ and covariances zero. The column vector, $\hat{\beta}$, given by C.4 now becomes B as in equation (5.3) and is used throughout the entire next block.

B. The ninth observation is taken at treatment combination abc with an observed response of 69 so the required data for equation (5.3) is $N = 8$, $n = 1$, $q = 8$, $m = 1$, $r_9^T = [1,1,1,1,-1,-1,-1,-1]$, $y_9 = 69$, $\hat{y}_9 = r_9^T B = 60$, and B as given in (C.4).

This results in

$$B^* = \begin{bmatrix} 57.5 \\ 11.5 \\ 15.5 \\ 7.5 \\ 18.5 \\ 0 \\ 2.0 \\ 11.5 \end{bmatrix} \qquad (C.7)$$

so the fitted model after nine observations is

$$y = 57.5 + 11.5A + 15.5B + 7.5C + 18.5D + 2.0F + 11.5\ BD \qquad (C.8)$$

and the variance of the coefficients in (C.8) as obtained from (5.4) is $15/128\ \sigma^2$ but the covariances are no longer zero because the columns of the design matrix are no longer orthogonal. The covariances can be obtained from the approximate entry in the variance-covariance matrix, $(x'x)^{-1}$. This procedure would be continued through the 16^{th} observation. The B^* obtained after applying (5.3) to the 16th observation would become the values of B used throughout the third block of eight observations.

BIBLIOGRAPHY

1. Addelman, S., "Irregular Fractions of the 2^n Factorial Experiments," Technometrics, Vol. 3, 1961, 479-496.

2. Anderson, W. T., "A Modification of the Sequential Probability Ratio Test to Reduce Sample Size," Annals of Mathematical Statistics, Vol. 19, 1948, 326-339.

3. Armitage, P., "Sequential Analysis with More Than Two Alternative Hypotheses and Its Relation to Discriminant Function Analysis," Journal of the Royal Statistical Society, Series B, Vol. 12, 1950, 137-144.

4. Armitage, P., "Restricted Sequential Procedures," Biometrika, Vol. 44, 1957, 9-26.

5. Army Regulation 10-4, Organization and Functions of OTEA, December 1974.

6. Army Regulation 71-3, Research and Development and Force Development Testing, November 1974.

7. Army Regulation 1000-1, Basic Policies for Systems Acquisition by the Department of the Army, November 1974.

8. Atkinson, A. C. and Hunter, W. G., "The Design of Experiments for Parameter Estimation," Technometrics, Vol. 10, 1968, 271-289.

9. Baker, A. G., "Properties of Some Tests in Sequential Analysis," Biometrika, Vol. 37, 1950, 334-346.

10. Bechhofer, R. E., Kiefer, J. and Sobel, M., Sequential Identification and Ranking Procedures, University of Chicago Press, Chicago, 1968.

11. Blot, W. J. and Meeter, D. A., "Sequential Experimental Design Procedures," Journal of the American Statistical Association, Vol. 68, 1973, 586-583.

12. Box, M., "Some Experiences with a Non-linear Experimental Design Criteria," Technometrics, Vol. 12, 1970, 569-589.

13. Box, M., "Simplified Experimental Design," Technometrics, Vol. 13, 1971, 19-31.

14. Burkholder, D. and Wysman, R., "Optimum Properties and Admissibility of Sequential Tests," Annals of Mathematical Statistics, Vol. 34, 1963, 1-17.

15. Chernoff, H., "Sequential Design of Experiments," Annals of Mathematical Statistics, Vol. 30, 1959, 755-770.

16. Daniel, C., Applications of Statistics to Industrial Experimentation, John Wiley and Sons, New York, 1976.

17. David, H. T. and Kruskal, W. H., "The WAGR Sequential t-test Reaches a Decision with Probability One," Annals of Mathematical Statistics, Vol. 27, 1956, 797-805.

18. Department of Defense Directive 5000-1, Acquisition of Major Defense Systems, U.S. Government Printing Office, Washington, D.C., September 1974.

19. Department of Defense Directive 5000-2, The Decision Coordinating Paper (DCP and the Defense Systems Acquisition Review Council (DSARC), U.S. Government Printing Office, Washington, D.C., February 1974.

20. Ehrenfeld, S., "On the Efficiency of Experimental Designs," Annals of Mathematical Statistics, Vol. 26, 1955, 247-255.

21. Fry, R. E., "Finding New Fractions of Factorial Experimental Designs," Technometrics, Vol. 3, 1961, 359-370.

22. Ghosh, B. K., Sequential Tests of Statistical Hypotheses, Addison Wesley, Reading, Mass., 1970.

23. Govindarajulu, Z., Sequential Statistical Procedures, Academic Press, New York, 1975.

24. Hall, W. J., "Some Sequential Analogs of Stein's Two-Stage Test," Biometrika, Vol. 49, 1962, 367-378.

25. Hicks, C. R., Fundamental Concepts in the Design of Experiments, Holt, Rinehart and Winston, New York, 1973.

26. Hines, W. W. and Montgomery, D. C., Probability and Statistics in Engineering and Management Science, The Ronald Press Company, New York, 1972.

27. Hoeffding, W., "Lower Bounds for the Expected Sample Size and the Average Risk of a Sequential Procedure," Annals of Mathematical Statistics, Vol. 31, 1960, 352-368.

28. Hoel, P., Port, S. and Stone, C., Introduction to Statistical Theory, Houghton Mifflin Company, Boston, 1971.

29. Hunter, J. S. and Box, G. E. P., "The 2^{K-P} Fractional Factorial Designs, I," Technometrics, Vol. 3, 1961, 311-351.

30. Hunter, J. S. and Box, G. E. P., "The 2^{K-P} Fractional Factorial Designs, II," Technometrics, Vol. 3, 1961, 449-458.

31. Hunter, J. S., "Sequential Factorial Estimation," Technometrics, Vol. 6, 1964, 41-55.

32. Jackson, J. E. and Bradley, R., "Sequential X^2 and T^2 Tests and Their Application to an Acceptance Sampling Problem," Technometrics, Vol. 3, 1961, 519-534.

33. John, P., Statistical Design and Analysis of Experiments, The Macmillan Company, New York, 1971.

34. Johnson, N. L., "Sequential Analysis: A Survey," Journal of the Royal Statistical Society, Series A, Vol. 124, 1961, 372-411.

35. Kemp, K. W., "Formulae for Calculating the Operating Characteristic and the Average Sample Number of Some Sequential Tests," Journal of the Royal Statistical Society, Series B, Vol. 20, 1958, 379-386.

36. Kempthorne, O., The Design and Analysis of Experiments, John Wiley and Sons, New York, 1952.

37. Kempthorne, O., "The Efficiency Factor of an Incomplete Block Design," Annals of Mathematical Statistics, Vol. 27, 1956, 846-849.

38. Kiefer, J. and Weiss, L., "Some Properties of Generalized Sequential Probability Ratio Tests," Annals of Mathematical Statistics, Vol. 28, 1957, 57-75.

39. Kshirsagar, A., "A Note on Incomplete Block Designs," Annals of Mathematical Statistics, Vol. 29, 1958, 907-910.

40. Lehmann, E. L., Testing Statistical Hypotheses, John Wiley and Sons, Inc., New York, 1959.

41. Matthes, T. K., "On the Optimality of Sequential Probability Ratio Tests," Annals of Mathematical Statistics, Vol. 34, 1963, 18-21.

42. Montgomery, D. C., Design and Analysis of Experiments, John Wiley and Sons, New York, 1976.

43. Mote, V. L., "On a Minimax Property of a Balanced Incomplete Block Design," Annals of Mathematical Statistics, Vol. 29, 1958, 910-913.

44. Neuhardt, J. B. and Bradley, H. E., "On the Selection of Multi-factor Experimental Arrangements with Resource Constraints," Journal of the American Statistical Association, Vol. 66, 1971, 618-621.

45. Neuhardt, J. B. and Bradley, H. E., "Computational Results in Selecting Multifactor Experimental Arrangements," Journal of the American Statistical Association, Vol. 68, 1973, 608-611.

46. Page, E. S., "An Improvement to Wald's Approximation for Some Properties of Sequential Tests," Journal of the Royal Statistical Society, Series B, Vol. 16, 1954, 136-139.

47. Read, C. B., "The Partial Sequential Probability Ratio Test," Journal of the American Statistical Association, Vol. 66, 1971, 646-650.

48. Russ, S. W., "A Cost Optimal Approach to Selection of Experimental Designs for Operational Testing Under Conditions of Constrained Sample Size," Georgia Tech Master's Thesis, 1976.

49. Sedransk, J., "Designing Some Multi-factor Analytical Studies," Journal of the American Statistical Association, Vol. 62, 1967, 1121-1139.

50. Sobel, M. and Wald, A., "A Sequential Decision Procedure for Choosing One of Three Hypotheses Concerning the Unknown Mean of a Normal Distribution," Annals of Mathematical Statistics, Vol. 20, 1949, 502-522.

51. U.S. Army Operational Test and Evaluation Agency, Operational Test and Evaluation Methodology Guide, May 1976.

52. Wald, A., Sequential Analysis, John Wiley and Sons, New York, 1947.

53. Wald, A. and Wolfowitz, J., "Optimum Character of the Sequential Probability Ratio Test," Annals of Mathematical Statistics, Vol. 19, 1948, 326-339.

54. Wald, A. and Wolfowitz, J., "Bayes Solution of Sequential Decision Problems," Annals of Mathematical Statistics, Vol. 21, 1950, 82-89.

55. Weiss, L., "Testing One Simple Hypothesis Against Another," Annals of Mathematical Statistics, Vol. 24, 1953, 273-281.

56. Wetherill, G. B., Sequential Methods in Statistics, John Wiley and Sons, Inc., New York, 1975.

57. Yates, F., Design and Analysis of Factorial Experiments, Imperial Bureau of Soil Sciences, London, 1937.

58. Youden, W. J., "Partial Confounding in Fractional Replication," Technometrics, Vol. 3, 1961, 353-358.